

Beyond the Light Cone: Causal Influence Density and the Entropy Arrow as a Third Axis for Minkowski Diagrams

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

February 9, 2026

Minkowski diagrams are among the most powerful visual tools in relativity: with a single 2D plot, they show why simultaneity is relative, why the speed of light is a causal limit, and how proper time organizes motion. Yet their strength is also their limitation. A standard Minkowski diagram is a geometry of possibility: it tells us what can, in principle, influence what. It does not, by itself, tell us what does influence what, nor why time in lived physics feels directionally loaded. The light cone is symmetric; the arrow is not. This paper proposes a speculative but concrete extension: keep the 2D Minkowski base plane (ct, x) and add a third, non-spatial axis representing a scalar field on spacetime, $\Phi(ct, x)$, interpreted as causal accessibility, information-flow capacity, or entropy-arrow strength. Instead of adding another spatial coordinate, the diagram is “lifted” into a causal topography where regions of high Φ correspond to events with large accessible causal pasts (relative to a record-bearing domain), or to zones of irreversible macro-amplification where entropy production is concentrated. The result is a hybrid visualization: Minkowski geometry supplies the invariant causal skeleton, while Φ supplies a physically motivated asymmetry emerging from matter, records, and boundary conditions.

Keywords: Minkowski diagram, spacetime visualization, causal influence density, information flow, entropy arrow, thermodynamic time, causal accessibility, light cone geometry, Lorentz invariance, proper time, records and memory, irreversible processes.

A collaboration with GPT-5.2-Thinking. 59 pages. CC4.0.

Outline

1. Introduction: From Causal Possibility to Causal Topography
 - 1.1 Why Minkowski diagrams remain uniquely pedagogical
 - 1.2 The missing ingredient: directionality, records, and “what actually happens”
 - 1.3 Proposal in one sentence: add a non-spatial height field $\Phi(ct, x)$
 - 1.4 What this paper is and is not claiming
2. The Standard Minkowski Diagram as a Causal Skeleton
 - 2.1 Events, worldlines, and the light cone as a boundary of influence
 - 2.2 Timelike, lightlike, spacelike separation and what “causal” means
 - 2.3 Invariant interval and why the diagram is symmetric under time reversal
 - 2.4 Compression of space and the interpretive cost of 2D
3. Why the Arrow of Time Is Not in Minkowski Geometry
 - 3.1 Time-reversal symmetry of the equations versus asymmetric histories
 - 3.2 Boundary conditions, low-entropy past, and the emergence of direction
 - 3.3 Records, memory, and amplification as physical asymmetry mechanisms
 - 3.4 A diagrammatic goal: show arrow-structure without rewriting relativity
4. Defining the Third Axis: $\Phi(ct, x)$ as a Field on Spacetime
 - 4.1 Coordinate versus field: why “not spatial” matters
 - 4.2 Minimal mathematical requirements for Φ (scalar, nonnegative, interpretable)
 - 4.3 Three interpretations of Φ : accessibility, influence, and entropy production
 - 4.4 Choosing observables and acknowledging observer-dependence
5. Model A: Causal Accessibility as Past-Cone Overlap with a Record Domain
 - 5.1 The record-bearing domain D as a world-tube in spacetime
 - 5.2 $\Phi(p) = \log(1 + \text{measure}(\text{PastCone}(p) \cap D))$ and “bit-like” scaling
 - 5.3 Geometric intuition: why Φ grows, saturates, and forms domes/ridges
 - 5.4 What this model says about “how much of the world could have informed me”
6. Model B: Influence Propagation and Retarded-Front Topographies
 - 6.1 From possibility to propagation: kernels, fronts, and wakes
 - 6.2 Visualizing sources: ridges on future cones rather than abstract boundaries
 - 6.3 Superposition of influences and the emergence of causal “weather maps”
 - 6.4 How this differs from field theory diagrams and why it remains Minkowski-based
7. Model C: Entropy Production and Irreversibility Landscapes
 - 7.1 Φ as coarse-grained entropy S or local production rate σ
 - 7.2 Entropy cliffs, ramps, and localized bumps as macro-amplification events

- 7.3 The measurement analogy: microscopic symmetry, macroscopic irreversibility
- 7.4 Why this can visualize an arrow without claiming time itself is generated
- 8. Symmetry, Invariance, and Interpretation Constraints
 - 8.1 Can Φ be Lorentz-scalar, and when should it not be?
 - 8.2 Frame dependence as a feature: observers, records, and operational meaning
 - 8.3 Avoiding category errors: Minkowski geometry versus added physical structure
 - 8.4 What “objectivity” could mean for information and entropy in this diagram
- 9. Worked Example: The 3D Lifted Diagram and How to Read It
 - 9.1 Diagram components: base plane, light cone lines, record world-tube boundaries
 - 9.2 Interpreting the surface: low- Φ deserts, high- Φ plateaus, entropy bumps
 - 9.3 Diagnostic readings: where information is available, where arrow-strength concentrates
 - 9.4 Extensions: multiple domains, competing observers, and moving record-tubes
- 10. Pedagogy and Research Utility: What This Adds Beyond Standard Diagrams
 - 10.1 Teaching causal structure plus arrow-structure in one unified picture
 - 10.2 Bridging relativity and thermodynamics without collapsing one into the other
 - 10.3 Visualization as hypothesis generator: where to expect records, fragility, or amplification
 - 10.4 Limits, failure modes, and what would count as a better definition of Φ
- 11. Open Problems and Future Directions
 - 11.1 Generalizing to 3+1: from surfaces to volumes and isosurfaces of Φ
 - 11.2 Curved spacetime: replacing cones with causal diamonds and horizons
 - 11.3 Connections to causal set ideas, coarse-graining, and computational mechanics
 - 11.4 Toward a standardized “causal topography” toolkit
- 12. Conclusion: A New Visual Grammar for Causality and the Arrow
 - 12.1 Minkowski provides the skeleton; Φ provides lived asymmetry
 - 12.2 The arrow as a field over spacetime, not a coordinate of spacetime
 - 12.3 Why this may sharpen intuition about records, memory, and irreversibility
- References

1. Introduction: From Causal Possibility to Causal Topography

Relativity's most durable insights are not only mathematical; they are visual. A Minkowski diagram takes ideas that otherwise feel like philosophical puzzles and turns them into geometry you can look at. With two axes and a few lines, it becomes obvious why “now” is not universal, why the speed of light is not merely a speed limit but a causal limit, and why proper time is the organizing clock of motion rather than an optional coordinate choice. For more than a century, this diagram has remained one of the best examples of physics pedagogy doing what it ought to do: converting abstraction into intuition without falsifying the underlying theory.

Yet the very elegance of the Minkowski diagram also reveals a boundary. It is a diagram of what is possible, not a diagram of what is typical, likely, recorded, amplified, or irreversible. It is a causal skeleton, not a lived landscape. It draws the light cone as a symmetric structure, and in doing so it tells the truth about the geometry of spacetime, but it also hides the truth about how real physical histories unfold inside that geometry. The diagram can tell you whether an influence can reach you; it cannot, by itself, tell you how much of the world did reach you, or why some influences become indelible records while others disperse into thermal noise. In the ordinary Minkowski picture, the future and the past are mirror images; in ordinary experience and ordinary thermodynamics, they are not.

This paper proposes a speculative but concrete visualization designed to sit on top of the Minkowski diagram rather than replace it. The proposal is to keep the Minkowski base plane (ct, x) intact, preserving its invariant causal structure, but to “lift” the plane into a third dimension that is explicitly not spatial. The new axis is a scalar field $\Phi(ct, x)$: a height function over events, intended to represent causal accessibility, information-flow capacity, or entropy-arrow strength. In this lifted picture, the light cone remains the boundary of causal reach, but the interior is no longer uniformly featureless. Instead, it becomes a topography: regions rise where accessible causal history is abundant or where irreversible amplification concentrates records, and regions remain low where causal access is sparse or where irreversibility is weak. The result is a diagram that makes a new claim only at the visualization level: that there is value in separating the causal skeleton of spacetime from the physically contingent “arrow content” supplied by matter, memory, boundary conditions, and coarse-graining.

1.1 Why Minkowski diagrams remain uniquely pedagogical

The Minkowski diagram compresses multiple non-intuitive consequences of Lorentz invariance into a single coherent geometry. This matters because many of relativity's core claims are not difficult in algebraic form but remain slippery in conceptual form. Algebra can tell you that simultaneity is relative; the diagram lets you see it. Algebra can tell you that proper time is maximized along inertial paths; the diagram lets you compare path lengths in the correct

pseudo-Euclidean sense. Algebra can tell you that superluminal signals would scramble causal order; the diagram makes the causal paradox immediate by showing how spacelike-separated events swap temporal order across frames.

Pedagogically, it does something rarer: it preserves the essential invariants while permitting multiple interpretations at once. You can talk about an event as a point, a particle as a worldline, a signal as a lightlike trajectory, and a “frame” as a choice of axes. You can also show that different frames are not disagreements about facts but different decompositions of the same invariant causal structure. In other words, the Minkowski diagram is a training device for intellectual humility: it forces a separation between what is invariant and what is coordinate-dependent.

It also has a second virtue: it makes causality visible without requiring metaphysical commitment. You do not need to settle debates about “what time really is” to use it. You simply accept the operational rule that influences propagate at or below c , and the diagram turns that into a geometric boundary condition. The diagram works because it is a map of constraint. It tells you what the universe will not permit, and that is often more pedagogically stable than telling you what the universe “is.”

1.2 The missing ingredient: directionality, records, and “what actually happens”

But precisely because the Minkowski diagram is a map of constraint, it is silent about the internal texture of allowed histories. Inside the light cone, almost everything is permitted by geometry. Yet not everything happens with equal likelihood, and not everything that happens leaves durable traces. Real physical systems do not merely evolve; they record, amplify, dissipate, and forget. Lived time has an arrow not because relativity inserts one, but because the world’s matter distribution, initial conditions, and coarse-grained dynamics select an asymmetry. A spilled drop of ink diffuses. A broken glass does not spontaneously reassemble. A macroscopic record, once amplified into many degrees of freedom, becomes robust; its time-reversed equivalent is possible in principle but vanishingly unlikely in practice.

This is not a minor add-on. It changes what it means to interpret causal statements. The bare Minkowski question is binary: can A influence B? The lived question is graded: how much of A’s past is available to B, how reliable is that information, and through what amplifying pathways does it become a record that can constrain future action? In the real world, causality is not only about signal speed; it is about record formation. A causal influence that leaves no record is still an influence in a microscopic sense, but it does not become part of the macroscopic narrative that agents inhabit. Conversely, once influence is recorded, it becomes a constraint on future worlds, not by violating relativity, but by being copied into many redundant degrees of freedom that make “forgetting” improbable.

In ordinary Minkowski space, the diagram is symmetric under time reversal: the past and future cones have the same shape. But in thermodynamic reality, the past is where records come from and the future is where records go. That statement is not a theorem of special relativity; it is a statement about how the universe has, in fact, arranged its boundary conditions and its irreversible pathways. If one wants a diagram that helps readers hold both truths at once, one needs a way to display the Minkowski constraint skeleton and the emergent arrow texture simultaneously.

1.3 Proposal in one sentence: add a non-spatial height field $\Phi(ct, x)$

The proposal is simple to state. Keep the standard 2D Minkowski diagram in the (ct, x) plane, but define a scalar field $\Phi(ct, x)$ that assigns to each event a nonnegative “height” representing one of the following closely related quantities: (i) causal accessibility to a record-bearing domain (how much causally available history is present), (ii) information-flow or influence capacity under a chosen propagation model (how much influence arrives or departs), or (iii) entropy-arrow strength (how much irreversible macro-amplification or entropy production occurs). Then plot the surface $z = \Phi(ct, x)$ above the Minkowski plane.

The effect is to turn a flat causal map into a causal landscape. The light cone lines remain the same; they still delimit possibility. But within the allowed region, Φ introduces a graded structure that is not geometrical in the strict Minkowski sense, but physical in the sense of matter, records, and irreversibility. Where Φ is high, the event sits atop a “ridge” of accessible history or irreversible amplification; where Φ is low, the event lies in a “desert” of sparse causal access or weak arrow content.

This is not an attempt to smuggle a new dimension into spacetime. It is the standard move of physics visualization: plotting a field over a manifold. The manifold is still Minkowski spacetime (reduced to 1+1 for drawing). Φ is extra structure, and the point is to see how extra structure can make time look one-way inside a geometry that, by itself, is two-way.

1.4 What this paper is and is not claiming

This paper is claiming that there is a useful conceptual separation between (a) the invariant causal skeleton provided by Minkowski geometry and (b) the emergent arrow content provided by physical processes that generate records and irreversibility. It is claiming that a visualization can help readers avoid a common confusion: the confusion of attributing the arrow of time to relativity itself, or conversely, the confusion of treating relativity as if it negates the arrow as a mere illusion. The arrow is not a coordinate artifact; it is a contingent field-like feature of how the universe populates spacetime with memory-bearing matter and dissipative dynamics.

This paper is not claiming that Φ is unique. Different definitions of “information,” “record,” and “entropy production” will generate different Φ fields. That is expected. Indeed, part of the value

of the approach is that it makes those choices explicit rather than implicit. It is also not claiming that Φ must be Lorentz invariant in all interpretations; some Φ definitions are properly frame-dependent because information and records are operationally tied to observers, apparatus, and coarse-graining choices. Where Lorentz-scalar constructions are available, they can be emphasized; where they are not, the dependence can be treated as a feature revealing the operational nature of information.

Most importantly, this paper is not claiming a new physical theory of time. It is offering a diagrammatic grammar: a way to place thermodynamic and informational asymmetry into the same visual field as relativistic causality without collapsing one into the other. Minkowski remains Minkowski. The light cone remains the light cone. The new element is a plotted height: a reminder that inside the cone, physics is not only about what is permitted but about what is amplified into history.

If the paper succeeds, the reader should come away with a single combined intuition: relativity supplies the boundaries of causal reach, while thermodynamics and information supply the terrain of irreversible narrative. The arrow is not a new axis of spacetime. It is a field that lives on spacetime, sculpted by boundary conditions, matter, and the dynamics of record formation.

2. The Standard Minkowski Diagram as a Causal Skeleton

Before we add any “height field” or arrow-content, we need to treat the standard Minkowski diagram with the respect it deserves: it is one of the cleanest visual encodings of causal structure ever invented. In its simplest form, it is a 2D plot with one axis representing time (often ct so that both axes share length units) and one axis representing a single spatial coordinate x . Every point in the plane is an event, and every curve is a possible history of something extended in time: a particle, a signal, a clock, a spaceship, a laboratory. What makes the diagram powerful is not the axes alone but the geometry induced by the speed-of-light constraint. The diagram does not merely depict motion; it depicts what kinds of motion and influence are physically admissible.

Calling it a “causal skeleton” is not rhetorical. The Minkowski diagram strips away almost everything except the minimal, invariant structure that must be respected by any physical process: no influence that carries information propagates faster than light, and Lorentz transformations preserve this causal structure. That skeleton is not the whole story of physics, but it is the frame upon which many stories must be hung. Once you draw the light cone, you have already imposed a fundamental ordering constraint on the universe: which events can possibly be connected by cause and effect.

2.1 Events, worldlines, and the light cone as a boundary of influence

In a Minkowski diagram, an event is represented by a point. This is conceptually important: an “event” is not a thing, not an object, not a location in space, but a specific occurrence at a specific place and time. A flash of light, a collision, the ticking of a clock, the emission of a photon, the pressing of a button. Each can be idealized as an event.

A worldline is a curve traced by an object through spacetime. A stationary object in the chosen frame has a worldline that is vertical (in the ct direction): x is constant while time advances. An object moving at constant velocity has a straight worldline with some slope: x changes as time advances. The steeper the worldline, the slower the spatial motion; the closer it tilts toward the 45-degree line, the faster the spatial motion. This “slope intuition” is one of the diagram’s pedagogical gifts: you can see velocity as geometry.

The light cone is drawn by taking an event (often the origin) and plotting the paths that light could take leaving that event. In units where $c = 1$ (or by using ct on the vertical axis), light rays are lines at 45 degrees: $x = \pm t$. These lines form the boundaries of the future and past light cones. Everything inside the future light cone consists of events that can be reached from the origin by influences traveling at speeds less than or equal to c . Everything inside the past light cone consists of events that could have influenced the origin. The regions outside both cones are spacelike separated from the origin and cannot be connected by any causal influence constrained by c .

This is why the light cone is a boundary of influence rather than merely a depiction of “light.” It is the geometry of maximum causal reach. Even if no light is present, the cone expresses the ultimate constraint on any signal or influence that carries information. The cone is therefore not a special feature of electromagnetism; it is the organizing boundary of relativistic causality.

2.2 Timelike, lightlike, spacelike separation and what “causal” means

With the light cone in place, the diagram naturally partitions all separations between events into three classes.

Timelike separation occurs when one event lies inside the other’s light cone. If event B lies in the future light cone of event A, then there exists some subluminal worldline that can connect A to B. In ordinary language, A can cause B in principle, because there is enough time for a slower-than-light influence to travel between them. In the Minkowski diagram, timelike separated events can be assigned a consistent order for all inertial observers: if B is in A’s future light cone, no inertial frame will reverse that causal order. This invariance of causal order is one of the deepest reasons relativity is not “mere perspective.” Perspectives differ about simultaneity for spacelike separated events, but they do not differ about what can be causally downstream.

Lightlike separation occurs when event B lies exactly on the light cone of event A. Then the only possible connection between them is at the speed of light. This is the boundary case: there is just enough time for a light signal to connect them, but no slower influence could make it. Lightlike paths form the edges of causal possibility. They are the “limit curves” between what can be connected and what cannot.

Spacelike separation occurs when event B lies outside the light cone of event A. In this case, no influence constrained by c can connect them. This is where relativity rewires the intuitive notion of “now.” Spacelike separated events do not have an invariant temporal order. Depending on the inertial frame, one may be judged to happen before the other, or after, or simultaneously. This is not paradoxical; it is a geometric consequence of how different frames slice spacetime into “space at a given time.” What remains invariant is not the temporal order but the impossibility of causal connection.

From these distinctions, “causal” acquires a precise meaning in this diagrammatic setting: a causal relationship is only possible when events are timelike or lightlike separated. Spacelike relationships are correlations without causal linkage, unless one introduces superluminal influences (which would typically generate causal loops and contradictions in standard relativistic frameworks). The Minkowski diagram thus enforces a discipline: it forces you to separate correlation, simultaneity, and influence into categories that cannot be casually blended.

2.3 Invariant interval and why the diagram is symmetric under time reversal

The Minkowski diagram is not Euclidean geometry drawn on graph paper. It is a representation of a pseudo-Euclidean structure where the invariant “distance” between events is not $x^2 + (ct)^2$ but $(ct)^2 - x^2$ (in 1+1 dimensions). This quantity, the invariant interval, is what all inertial observers agree on, even though they disagree about the separate components ct and x .

The interval’s sign encodes the three separation types. If $(ct)^2 - x^2$ is positive, the separation is timelike. If it is zero, the separation is lightlike. If it is negative, the separation is spacelike. This single expression therefore contains the causal classification of event separations. It is not merely a formula; it is the algebraic backbone of the light cone itself.

Now consider symmetry under time reversal. In the bare geometry, if you flip t to $-t$, the interval remains unchanged because t appears squared. The light cone is symmetric: the future cone and the past cone have identical shape. If you run the diagram backward, the causal structure is the same. This is the central fact that motivates the later addition of Φ . Minkowski geometry alone does not prefer the future over the past. It is a skeleton of constraint that is indifferent to narrative direction.

This does not mean physics is time-reversal symmetric in every sense. It means the geometry that constrains causal reach is symmetric. Any arrow of time must therefore come from something not contained in the Minkowski metric alone: boundary conditions, statistical mechanics, coarse-graining, dissipative dynamics, cosmological initial low entropy, or other asymmetries introduced by the matter content and history of the universe. The diagram, as usually drawn, has no place to display that asymmetry. It is honest, but it is silent.

That silence is not a flaw in Minkowski space; it is a reminder that causal order and thermodynamic direction are different ideas. Minkowski gives you “which events can influence which.” Thermodynamics gives you “which macrostates are typical, which transitions are irreversible, and which traces become records.” The latter is not written into the metric.

2.4 Compression of space and the interpretive cost of 2D

A standard Minkowski diagram collapses three spatial dimensions into one. This is not merely a technical simplification; it is a conceptual choice with consequences. In the physical world, causal influence spreads outward in all spatial directions. In a 1+1 diagram, that spreading is represented by two lines, $x = \pm t$, forming a wedge. In 3+1 dimensions, the analog is a light cone whose cross-sections are spheres expanding at speed c . The 2D diagram captures the causal boundary correctly, but it hides the geometry of spatial spreading that is often essential for thinking about information, dissipation, and records.

This compression has several interpretive costs.

First, it can exaggerate the sense that “everything happens along a line.” In 1+1, influence is constrained to the single spatial axis drawn, whereas in reality influence can bypass, scatter, diffuse, and disperse through a vast transverse space. Many thermodynamic processes are profoundly spatial in their mechanism: diffusion, mixing, turbulence, radiation into many modes. Compressing to one spatial dimension can make irreversibility look like a property of time alone rather than a property of how influence distributes into many degrees of freedom.

Second, it can hide the role of redundancy and record multiplication. A record is robust not merely because an event occurred, but because the information about it becomes copied into many independent degrees of freedom: photons scattering off an object, thermal excitations, macroscopic marks, neural states, printed text, digital storage. In 3D space, the number of channels for such copying is enormous. In 1D, it is artificially limited. So a 2D Minkowski diagram can teach causality beautifully while under-teaching how causality becomes memory.

Third, it can obscure the difference between “can influence” and “will influence with any appreciable strength.” In higher dimensions, influence can attenuate with distance and spread over expanding wavefronts. In 1+1, the geometric intuition for attenuation is less natural. That

matters for any attempt to plot an information or entropy-related field Φ : the shape of Φ depends strongly on whether causal influence spreads into a line, a plane, or a volume.

These costs do not make 2D Minkowski diagrams less valuable; they simply define the boundary of what the diagram is designed to do. It is a skeleton, not a full organism. It gives you the invariant causal partition of spacetime but not the density of influence, not the robustness of records, and not the emergent arrow that comes from the world's occupation of spacetime.

This is the point at which our proposed extension becomes motivated. If we want to keep the pedagogical power of the Minkowski diagram while also giving readers a place to “see” asymmetry, we need a representation that respects the causal skeleton but permits additional structure to ride on top of it. The next step is to formalize that additional structure as a non-spatial field $\Phi(ct, x)$ and to argue for principled ways of interpreting it.

3. Why the Arrow of Time Is Not in Minkowski Geometry

A persistent temptation in modern thought is to treat the arrow of time as something “built into” the structure of spacetime itself. Minkowski diagrams, with their clean cones and invariant intervals, can accidentally reinforce this temptation: the diagram looks like a cosmic rulebook, so one expects it to contain everything important about time. But in fact the opposite is true. Minkowski geometry is deliberately minimal. It specifies causal constraints and invariant structure, but it does not encode the thermodynamic, informational, and narrative features that make time feel directionally loaded. This is not a defect; it is the profound clarity of the theory. The arrow of time is not a geometric necessity. It is a physical contingency. Minkowski gives the stage; the arrow is in the play.

The claim of this section is straightforward: if you want to explain why time has a direction in the world we inhabit, you must add something beyond Minkowski geometry—something like boundary conditions, coarse-graining, dissipative dynamics, or the physics of records. That “something” can be made explicit without altering the causal skeleton. The argument matters for this paper because the proposed 3D “causal topography” is exactly an attempt to display what Minkowski geometry omits, while keeping Minkowski geometry intact.

3.1 Time-reversal symmetry of the equations versus asymmetric histories

Many of the fundamental dynamical equations used in physics are time-reversal symmetric in the following sense: if you take a physically allowed micro-trajectory and reverse the sign of time (and make the necessary associated reversals, such as flipping velocities or momenta), the reversed trajectory is also consistent with the equations. This is not universally true in every

corner of physics, but it is broadly true enough that it forces a conceptual separation: the laws, at the microscopic level, often do not prefer a direction of time.

Yet we encounter time as directionality. We remember the past but not the future. We see mixing, diffusion, and decay, but not their spontaneous reverse. We observe that macroscopic systems move from special to typical states. The puzzle is not that this happens—statistical mechanics explains it in a robust way—but that the asymmetry is so pervasive that it can feel like a built-in property of time itself.

The key reconciliation is that symmetry of laws does not imply symmetry of realized histories. A symmetric rule can still generate asymmetric outcomes if the initial state is not symmetric, or if the relevant description is coarse-grained and includes a statistical weighting over microstates. In practice, the world's microstate at any moment is not a generic random point in phase space; it is constrained by how it got there. As soon as you admit “how it got there,” you are already dealing with history. And history, in our universe, appears to begin in an extraordinarily special condition relative to what is typical at later times.

Minkowski geometry mirrors this distinction. The causal structure of Minkowski spacetime is symmetric under time reversal: the past light cone and the future light cone have the same shape. Nothing in the geometry tells you which direction is “forward.” A Minkowski diagram can be read with time pointing up or down; it does not protest. If you draw the same diagram and reverse the direction of the time axis, the causal classification of separations remains unchanged. In other words: Minkowski spacetime has no intrinsic arrow, and it is not supposed to.

3.2 Boundary conditions, low-entropy past, and the emergence of direction

If the arrow is not in the geometry and not in the micro-laws alone, where does it come from? The most widely accepted answer is that it emerges from boundary conditions—especially the condition that the universe (or any isolated system under discussion) began in a remarkably low-entropy macrostate compared with what is typical.

This is often stated too casually, as if “low entropy in the past” were just a slogan. But it is more precise than that. In a coarse-grained description, a macrostate corresponds to a vast number of microstates. High-entropy macrostates correspond to overwhelmingly larger volumes of microstate space than low-entropy ones. If you select a microstate at random from a high-entropy macrostate, it is almost certainly “typical,” and its future evolution is overwhelmingly likely to remain typical. If you condition instead on an atypical, low-entropy macrostate, the evolution will almost certainly move toward more typical macrostates simply because there are vastly more of them. The arrow appears as a probability gradient induced by the combinatorics of microstates, but that gradient requires a special boundary condition to get started.

Minkowski geometry does not contain that boundary condition. The geometry can host a universe with a low-entropy beginning, or one with a low-entropy ending, or one with symmetric boundary constraints at both ends, or one with no such constraint. The geometry itself does not choose. The arrow is therefore not a theorem of relativity but a contingent fact about which histories are realized.

This matters for visualization. A Minkowski diagram tells you what causal influences are possible, but it does not tell you whether the past is “special” in the thermodynamic sense. If we want a diagram that makes the arrow visible, we need a way to represent boundary conditions and their downstream consequences without pretending that they are part of Minkowski structure. This is exactly why a plotted field Φ is conceptually appropriate: Φ can encode “how arrow-loaded” the local region of spacetime is, given particular boundary conditions and matter distributions, without confusing that encoding with the metric itself.

3.3 Records, memory, and amplification as physical asymmetry mechanisms

Thermodynamics is sometimes presented as if it were about “disorder increasing.” That is a poor metaphor for our purposes. A better framing for this paper is that the arrow is about records: which events can leave traces that persist and become constraints on future dynamics. In lived physics, the past is not only what happened; it is what is recorded. The present is not only a moment; it is a moment saturated with inherited records. The future is not only what will happen; it is what is not yet recorded, not yet redundantly copied into the environment.

A record is not magic. It is a physical configuration that correlates with some earlier event and is stabilized by amplification. Consider a simple chain: a microscopic event triggers a mesoscopic change, which triggers a macroscopic effect, which then scatters into many environmental degrees of freedom. The key step is multiplication. Once information is copied into many independent channels, erasing it requires a coordinated reversal across all those channels—a fantastically improbable process. The arrow is therefore not just about energy dispersing; it is about correlations becoming redundantly encoded in a way that makes reversal practically impossible.

This is why “memory” is not an abstract psychological phenomenon in this context. Memory is a thermodynamic structure. A brain’s memory is physical, and so is a camera’s memory, and so is a rock’s “memory” if you define memory broadly enough: a track in a substrate, a thermal scar, a chemical gradient, a pattern of photons in flight. The world is filled with such records, and they overwhelmingly point one way: from earlier to later. The fundamental reason is not that Minkowski spacetime prefers that direction, but that physical amplification processes do.

If a third-axis field Φ is meant to visualize arrow-content, then record formation is one of its best candidates. A region of spacetime with intense amplification—where microscopic

differences get blown up into macroscopic, redundant records—should register as a “high Φ ” region. Conversely, a region where interactions are weak, where correlations do not get copied, where noise dominates and records are fragile, should register as low Φ . In this sense, Φ can be read as “record potential” or “record density” as much as “entropy arrow strength.” The two are tightly linked: the creation of robust records is a kind of entropy export into environmental degrees of freedom. The record becomes stable precisely because the environment has taken on the burden of redundancy.

This is also where observer-dependence enters in a controlled way. What counts as a “record” depends on coarse-graining and on what variables are treated as meaningful. But the underlying physical mechanism—amplification into many degrees of freedom—does not depend on philosophical taste. It is a fact about open systems and dissipation. So Φ can be defined in ways that are operationally anchored even when they are not uniquely determined.

3.4 A diagrammatic goal: show arrow-structure without rewriting relativity

At this point the aim of the visualization should be stated as a discipline: we want a diagram that shows arrow-structure, but we do not want to commit the category error of rewriting relativity into thermodynamics or rewriting thermodynamics into geometry. The Minkowski diagram should remain what it is: a representation of causal constraints and invariant structure. The arrow-content should appear as an added layer, not as a deformation of the causal skeleton itself.

This distinction is crucial because it prevents two common confusions.

The first confusion is to say: “If the arrow is real, then spacetime must have a built-in direction.” That would be to mistake a contingent field property of the universe for a necessary property of the metric. The arrow can be real and robust while still being contingent on boundary conditions and the physics of records.

The second confusion is to say: “If Minkowski is symmetric, then the arrow is illusory.” That would be to mistake geometric symmetry for physical symmetry of realized histories. Minkowski symmetry under time reversal does not mean histories are time-reversal symmetric; it means the stage does not force the play to have one ending rather than another. The world’s actual play is strongly biased, and that bias can be modeled and visualized without implying that the stage itself is biased.

The diagrammatic goal, then, is to preserve the clean Minkowski partition—past cone, future cone, spacelike elsewhere—while adding a graded visualization that shows how the interior of those cones is not uniform in lived physics. Some regions are dense with accessible causal history because they sit downstream of record-bearing matter. Some regions are arrow-strong because irreversible macro-processes occur there. Some regions are arrow-weak because

correlations disperse without leaving robust traces. The light cone remains the boundary of possibility, but the surface Φ makes visible the unevenness of realized causality.

In summary, Minkowski geometry gives us the invariant causal skeleton. The arrow of time emerges from boundary conditions, statistical typicality, and the physical machinery of records and amplification. A lifted diagram with a non-spatial height field Φ is a proposal to hold both truths at once: causality as constraint, and arrow as contingent structure living on that constraint.

4. Defining the Third Axis: $\Phi(ct, x)$ as a Field on Spacetime

The central move of this paper is deliberately modest in physics and deliberately ambitious in visualization. We do not add a new coordinate to spacetime. We do not alter Minkowski geometry. We take the familiar 2D Minkowski diagram—events as points in the (ct, x) plane—and we add a third plotted axis that represents a scalar field defined on that plane. The result is not “3D spacetime” but a 2D spacetime slice with a height overlay: a causal surface rising and falling above the causal skeleton.

This distinction matters because it keeps categories clean. A coordinate is a label used to locate events. A field is a property assigned to events. Coordinates tell you where and when. Fields tell you what the world is like there: temperature, density, curvature, potential, entropy production, information accessibility, record strength. The third axis in this paper is explicitly of the second kind. It is not a dimension of spacetime; it is a dimension of description.

In this section we define what it means to choose such a Φ , what minimal constraints make it mathematically and interpretively sensible, and how different physical interpretations lead to different but related Φ constructions. We also address the unavoidable fact that quantities involving “information” and “records” often depend on coarse-graining and on observers. That dependence is not a fatal flaw; it is part of what makes the arrow of time a physical phenomenon rather than a purely geometric one.

4.1 Coordinate versus field: why “not spatial” matters

When people imagine “making a Minkowski diagram 3D,” the obvious instinct is to add another spatial axis, turning (ct, x) into (ct, x, y) or (ct, x, y, z) in some projection. That is useful for certain geometrical intuitions, but it does not solve the problem this paper targets. Our problem is not that we lack room for extra space. Our problem is that Minkowski geometry, even in full 3+1 form, does not display the arrow. Adding y does not create directionality. It merely gives the light cone its proper spherical cross-sections.

The missing ingredient is not spatial extent; it is causal texture. We want to depict how much of the causal past is accessible, how much influence arrives from a record-bearing domain, where irreversible amplification happens, and where “history becomes sticky.” Those are not coordinates. They are properties that events possess depending on the matter content, boundary conditions, and coarse-graining that define what counts as a record.

Saying “the third axis is not spatial” is therefore a philosophical and pedagogical safeguard. It prevents the reader from misreading the construction as a new claim about ontology—“spacetime has an extra dimension”—when the actual claim is methodological—“we can visualize a physically meaningful scalar over spacetime.” This is standard practice in physics: we plot a field over a domain to expose structure that is otherwise invisible. The domain here is a Minkowski slice; the field is Φ .

Finally, “not spatial” is important because it aligns with the goal of keeping relativity intact. If we add a coordinate, we must worry about how it transforms under Lorentz transformations and whether it changes the spacetime metric. If we add a field, we only need to worry about whether it is defined consistently and interpreted honestly. Minkowski remains the base. Φ is an overlay.

4.2 Minimal mathematical requirements for Φ (scalar, nonnegative, interpretable)

To be useful rather than decorative, Φ must satisfy a few minimal requirements. These are not deep axioms; they are practical constraints that keep the diagram readable and the interpretation stable.

First, Φ should be a scalar field on the chosen 1+1 domain. That means it assigns a single real number to each event (ct, x) in the diagram. Scalars are diagrammatically natural: “height” is an intuitive representation of scalar magnitude. Vector fields could be drawn too, but they demand arrows, streamlines, and additional conventions; they are harder to integrate cleanly with the light cone geometry without visual clutter. The scalar choice is the simplest bridge from Minkowski geometry to arrow-content.

Second, Φ should be nonnegative, or at least have a natural baseline such that “zero” corresponds to “no arrow-content” in the chosen sense. This is not strictly necessary mathematically, but it is useful pedagogically. If Φ is meant to represent accessibility, influence capacity, or entropy production, negative values are usually interpretively awkward. Nonnegativity makes “higher means more” unambiguous.

Third, Φ must be interpretable in a way that can be explained without circularity. The definition should not depend on the very thing it is supposed to visualize. For example, defining Φ as “how much the arrow points forward here” is empty. Defining Φ as a measure of past-cone overlap with a record-bearing domain, or as local entropy production rate, is concrete. The

viewer can then read the landscape and connect it to an operational story: how signals propagate, how records accumulate, where irreversibility is concentrated.

Fourth, Φ should be stable under modest changes of resolution and discretization. Any plotted surface is a numerical approximation, so the field should not be hypersensitive to arbitrary grid choices unless that sensitivity is itself meaningful. For example, a delta-function-like “event” can be smoothed into a localized bump without changing the conceptual message. This requirement pushes Φ toward being an integrated or coarse-grained quantity rather than a microscopic one.

Fifth, Φ should coexist peacefully with the causal skeleton. That means it should not contradict the light cone structure. If Φ is defined in terms of causal accessibility, it should respect causal ordering: it should not claim accessibility from outside the past cone. If Φ is defined in terms of influence propagation, it should not produce “influence ridges” beyond lightlike bounds unless the model explicitly assumes superluminal channels. In short: Φ is not allowed to smuggle in causal violations that Minkowski forbids, unless the paper is explicitly exploring that violation as a counterfactual.

These constraints are enough to make Φ a legitimate object: a scalar field with nonnegative magnitude, physical interpretation, and causal compatibility.

4.3 Three interpretations of Φ : accessibility, influence, and entropy production

The paper emphasizes three closely related interpretations of Φ . They are not mutually exclusive; they are different “lenses” through which arrow-content can be represented.

The first interpretation is causal accessibility. Here Φ measures how much causally available past is present relative to some domain D where information, records, or sources exist. One concrete form is:

$$\Phi(p) = \log(1 + \text{measure}(\text{PastCone}(p) \cap D))$$

In a 1+1 diagram, “measure” can be area. The idea is intuitive: the larger the overlap between the event’s past cone and the record-bearing domain, the more causal material could in principle have reached that event. The log compresses scale and invites an information-like reading: large overlap corresponds to many potential causal channels, but with diminishing returns in subjective significance.

The second interpretation is influence capacity or influence density. Here Φ measures how much influence arrives at or departs from an event under a chosen propagation model. The simplest story is a set of sources emitting signals that propagate within the future light cone. One then constructs Φ as an integral over sources weighted by a retarded kernel. In visualization terms, future cones stop being mere boundaries and become wavefront-like ridges:

regions that are “causally downstream” become elevated, and the surface shows influence as a kind of causal weather map.

This interpretation is valuable because it connects to intuitive physical ideas: attenuation with distance, spreading over modes, superposition of influences, and the difference between the possibility of influence and its effective strength.

The third interpretation is entropy-arrow strength, often expressed through entropy production. Here Φ is tied to irreversible macroprocesses: mixing, dissipation, measurement amplification, decoherence-like spreading into environmental degrees of freedom, chemical reactions that generate heat, and so on. Φ can be chosen as a coarse-grained entropy S of a local subsystem or, more sharply, as a local entropy production rate σ . In the diagram, irreversible processes become cliffs, ramps, or bumps—localized “arrow engines” where micro-differences are amplified into stable macroscopic records and where reversal is statistically suppressed.

This third interpretation is the most directly connected to the everyday arrow of time. It answers not only “how much past is accessible” but “where in spacetime irreversibility is being manufactured.” It also makes the record story tangible: irreversible entropy production is often the price paid for robust recording.

The three interpretations therefore form a conceptual triangle. Accessibility is about how much causal past can reach an event. Influence is about how much causal impact flows through an event. Entropy production is about where causal flow becomes locked into irreversible macroscopic constraints. In many realistic settings, these co-vary. The diagram can be built with one interpretation, then refined by adding another as a term in Φ , producing a composite field with layered meaning.

4.4 Choosing observables and acknowledging observer-dependence

The moment we speak of “information,” “records,” and “accessibility,” we must acknowledge that these are not purely geometric primitives. They depend on what we choose to count, what we choose to ignore, and what operational context defines “a record.”

This dependence shows up at several levels.

First, defining the record-bearing domain D is a modeling choice. D might be a world-tube representing “the laboratory.” It might be the worldlines of observers. It might be a region of high matter density. It might be the portion of spacetime within which signals are emitted or detected. Different choices correspond to different questions. If D represents “where I have instruments,” then Φ becomes observer-centric. If D represents “where matter exists,” Φ becomes more cosmological.

Second, “measure” is a choice. In 1+1, area is convenient, but in 3+1 the natural invariant measure is not simply Euclidean volume on a coordinate slice. One can weight by intensity, by energy flux, by number of degrees of freedom, by channel capacity, or by a decay kernel representing attenuation and noise. These are not arbitrary decorations; they are the operational bridge between abstract cones and real influence.

Third, entropy and entropy production are not absolute without specification of coarse-graining and subsystem. Entropy is a macro-variable, and its value depends on what distinctions you consider meaningful. A gas has higher entropy when you ignore microscopic details and track only macroscopic ones. A record has meaning only relative to a set of macroscopic variables that define what counts as “the same record.” Therefore any entropy-based Φ carries an explicit interpretive frame: what system is being coarse-grained, and how.

Fourth, some Φ constructions may be Lorentz-invariant scalars, while others are naturally tied to a chosen frame. For pedagogical purposes, frame dependence is not always a vice. In fact, it can be a virtue: it reinforces that information is often operationally defined in an observer’s rest frame because detection, memory, and instrumentation are built into matter that has a rest frame. A camera has a worldline; a brain has a worldline. The “arrow-content” relevant to that agent is inherently agent-relative in its operational definition, even if the underlying microphysics is Lorentz covariant.

The correct discipline is to state the observer dependence explicitly rather than deny it. The field Φ is not being offered as a new universal coordinate. It is being offered as a way to visualize how the causal skeleton is populated with physically meaningful asymmetries that emerge when you specify matter, records, and boundary conditions. The observer is part of that specification.

In practice, this section’s guiding principle can be stated as follows: Minkowski spacetime gives the invariant causal structure shared by all inertial observers; Φ gives the contingent arrow-content associated with a chosen physical context. By keeping these layers separate—skeleton below, field above—we gain clarity rather than lose it. The diagram becomes honest about what relativity guarantees and honest about what the world’s history supplies.

5. Model A: Causal Accessibility as Past-Cone Overlap with a Record Domain

Model A is the simplest version of the “lifted Minkowski” idea because it is built directly from the causal skeleton itself. It does not ask us to invent new dynamics, new propagation laws, or new thermodynamic machinery. It asks a single question that Minkowski geometry is already designed to answer in a binary way—what can influence what—and turns that binary question into a graded one—how much of a designated, record-bearing part of the world lies inside my

causal past. In other words, we keep the light cone, but we stop treating everything inside it as equally relevant.

The model begins with a choice: identify a domain D in spacetime that represents where “records,” “sources,” or “information-bearing structure” exists. Then, for any event $p = (ct, x)$ in the diagram, look at its past light cone and compute how much of that cone intersects D . The intersection is the portion of the record-bearing domain that could, in principle, have sent influences to p without violating relativity. The size of that intersection becomes the raw accessibility measure. A log transform turns it into a more human-scaled quantity that behaves like “bits” in the loose sense of diminishing returns: once you already have enormous accessible past, adding more does not change the qualitative story as much.

This model is intentionally modest. It does not claim that everything in $\text{PastCone}(p) \cap D$ actually influenced p , or that p “knows” any of it. It only claims that the overlap provides a useful proxy for potential causal access in a universe where the existence of records and signals is not uniform across spacetime.

5.1 The record-bearing domain D as a world-tube in spacetime

In a 1+1 Minkowski diagram, the most natural depiction of a persistent physical region is a world-tube: a strip extending through time over a range of spatial positions. Concretely, let D be a set of events defined by conditions such as:

$$D = \{ (ct, x) : 0 \leq t \leq T, |x| \leq a \}$$

This is a vertical band centered at $x = 0$, extending from the initial time $t = 0$ up to some cutoff T . Interpreting D is where the model becomes flexible.

One interpretation is instrumental: D is “the laboratory world-tube,” the region where detectors exist and where records are stored. In this reading, D is not the whole universe but the observational apparatus. Φ then measures how much of the lab’s past could have influenced a given event.

A second interpretation is environmental: D is “the matter-bearing region,” the portion of spacetime populated with degrees of freedom capable of carrying information. In an empty universe, light cones exist but there is little to carry records; in a matter-rich region, light cones intersect abundant structure.

A third interpretation is social or cognitive: D is “the memory-bearing domain,” including brains, documents, sensors, and communication channels. Then Φ can be read as the size of the potentially available causal archive at an event. This interpretation resonates with the theme of records and arrow-content: a world-tube of record-bearing systems is exactly what transforms bare causality into lived history.

The world-tube choice is pedagogically convenient because it is easy to draw and it makes the overlap geometry tractable. But it also has conceptual value. A world-tube formalizes the intuition that records are not everywhere. They occupy physical substrates. They persist along worldlines. They are localized in space and endure through time. D is the diagrammatic representation of “there is something here that can store, transmit, and preserve information.”

In later generalizations, D can be made more realistic: it can move (a slanted tube), fluctuate (a tube with variable width), fragment (multiple disconnected tubes), or decay (a tube with time-dependent weighting). But the simple vertical tube already captures the essential point: causal accessibility depends not only on the shape of the light cone but on where record-bearing structure actually lies.

5.2 $\Phi(p) = \log(1 + \text{measure}(\text{PastCone}(p) \cap D))$ and “bit-like” scaling

Given an event p , define its past light cone $\text{PastCone}(p)$ as the set of events that can influence p at or below the speed of light. In 1+1, if p has coordinates (t, x) , then an event (τ, ξ) is in the past cone if $\tau \leq t$ and $|x - \xi| \leq (t - \tau)$ (in units with $c = 1$). The intersection $\text{PastCone}(p) \cap D$ is therefore the subset of the record-bearing tube that lies inside this inequality.

Now we need a “measure” of that intersection. In a 1+1 diagram, the simplest measure is just area in the (t, x) plane. That area has a direct intuitive meaning: it counts the spacetime “extent” of record-bearing events that could, in principle, have influenced p . More sophisticated measures are possible—weighting by signal intensity, by density of degrees of freedom, by reliability of records—but area is a clean starting point.

Define:

$$A(p) = \text{measure}(\text{PastCone}(p) \cap D)$$

Then define the height field:

$$\Phi(p) = \log(1 + A(p))$$

The “1 +” ensures the log is well-defined even when A is zero. The log itself is not claiming literal Shannon information content. It is a scaling choice with a recognizable meaning: when potential causal access is small, increases matter a lot; when potential causal access is already huge, additional increases matter less. This is the same qualitative behavior that “bits” and “evidence accumulation” often exhibit in practical reasoning. It is an ergonomic transform: a way to map a rapidly growing geometric quantity into a range that can be plotted and compared without the surface becoming numerically dominated by the largest values.

One can also justify the log transform by analogy to multiplicative growth of available pathways. In many systems, accessible degrees of freedom scale roughly multiplicatively with time and

spatial extent; taking the log turns multiplicative scaling into additive scaling, which is easier to interpret as “effective capacity.” But the paper’s discipline is to keep the claim modest: log here is a visualization choice that yields a “bit-like” reading, not a proof that the universe is literally computing bits in this measure.

5.3 Geometric intuition: why Φ grows, saturates, and forms domes/ridges

Once D is fixed as a world-tube and $\text{PastCone}(p)$ expands as we move to later times, the qualitative shape of Φ is almost inevitable.

At very early times (near the lower boundary of the diagram), an event’s past cone is small. It reaches only a short time into the past and only a narrow range of x . Its intersection with D is therefore small. Φ is low. In the 3D diagram, the surface begins near zero.

As t increases, the past cone of p opens wider and deeper. It reaches farther back in time and spans a wider region in space. Therefore it captures more of the world-tube D . The overlap $A(p)$ grows, and Φ rises. The landscape climbs. This is the simplest arrow-like feature: later events have access to larger portions of the record domain’s history.

However, growth does not continue indefinitely in the same way if D is finite in time or finite in space.

If the world-tube has a finite time extent—if it starts at $t = 0$ and ends at some T —then for events well after T , the intersection cannot include times beyond T because D does not exist there. In that case $A(p)$ saturates in time: once your past cone already covers the entire time span of D , pushing to later times no longer adds new D -volume to the overlap. Φ then approaches a plateau rather than rising without bound. In the surface plot, this looks like a leveling off into a broad mesa.

Similarly, if D is finite in space—bounded by $|x| \leq a$ —then events far away in x have reduced overlap, because their past cones must reach farther back in time to include that spatial region. For a fixed t , there is a spatial window where the past cone intersects the tube; outside that window, overlap drops sharply. This creates domes and ridges centered on the tube.

The key intuitions can be summarized as follows.

First, Φ tends to be highest near the spatial center of D , because those events are most directly downstream of the record-bearing tube.

Second, Φ tends to increase with time until it saturates (if D has a finite past extent), because more of D becomes causally available.

Third, Φ tends to decrease as you move spatially away from D, because your past cone must “reach” across space to intersect the tube, and for fixed time there may not be enough causal room.

In the 3D figure produced earlier, these intuitions appear as a broad hill centered around $x = 0$, rising with t , with a flattening tendency at higher t and a falloff toward large $|x|$. If additional localized arrow-content is added (such as a bump representing entropy production), it sits on top of this baseline dome like a localized mountain on a plateau. But Model A itself already generates a meaningful causal landscape without any thermodynamic term.

This is why Model A is pedagogically useful: the shape of the surface is not arbitrary. It is a direct consequence of the light cone geometry intersecting a finite record domain. The reader can derive the shape by pure Minkowski reasoning and then interpret the result as “graded causal access.”

5.4 What this model says about “how much of the world could have informed me”

Model A is best read as a measure of potential causal archive. It answers, in an idealized way, the question: how much record-bearing spacetime lies in my causal past. The emphasis is “could,” not “did.” It is not reconstructing actual influence chains; it is bounding them.

In a universe where record-bearing structure is unevenly distributed—some regions dense with memory and communication channels, other regions empty or chaotic—this distinction matters. Two events may be equally deep in time in coordinate terms but have radically different causal accessibility because one sits downstream of a rich record domain while the other is causally isolated. Standard Minkowski diagrams treat both as simply “inside the cone.” Model A treats one as living on a high plateau of potential information and the other as living in a desert of potential information.

This shifts how we talk about causality. Instead of a binary “causal or not,” we gain a graded “causally saturated or causally sparse.” That grading is precisely what lived experience suggests. Human and instrumental access to the past is not uniform; it depends on where records exist, how signals propagate, and how robustly information is stored. The world does not merely have a past; it has an archived past, unevenly.

Model A also clarifies a deep philosophical point without having to argue it explicitly: much of what we call “the past” is operationally equivalent to “the set of records available now.”

Minkowski geometry alone says the past cone contains what could have influenced us. Model A says: within that cone, some regions are rich in record-bearing events and some are not. The arrow of time, as lived, is an arrow of accumulating accessible records.

Finally, Model A naturally invites extensions that sharpen its realism. One can weight D by record density (not all spacetime points in the tube are equally recordful). One can weight past-cone contributions by attenuation and noise (not all causal channels are equally reliable). One can add decay of records over time (older records become less accessible). One can introduce moving observers and moving tubes (relativity of access). Each extension preserves the core discipline: the light cone remains the boundary, but the interior is given texture by a physically motivated field.

In this way, Model A provides a baseline “causal topography” that is purely Minkowski in its causal logic, but already rich enough to make the reader feel what Minkowski diagrams do not normally show: that causal possibility is not the same thing as causal accessibility, and accessibility is one of the simplest ways the arrow of time becomes visible without rewriting relativity.

6. Model B: Influence Propagation and Retarded-Front Topographies

Model A treated the light cone as a geometric container and asked how much of a record-bearing domain lies inside it. That is a measure of potential access. Model B shifts the emphasis from access to action. Instead of asking “how much could have influenced this event,” we ask “how much influence is arriving here, or leaving here, given actual sources and a propagation rule.” This step matters because arrow-content is not only about what the past could contain, but about what the world is actively doing: emitting, transmitting, attenuating, scattering, and imprinting.

The Minkowski diagram already knows the key rule for propagation: causal influence travels within the light cone. But it does not specify how influence is distributed within that cone, how it decays, how it accumulates, or how many sources contribute. Model B supplies exactly that missing layer by introducing a retarded propagation model. “Retarded” means future-directed: the influence at an event depends on sources in its past light cone (or, equivalently, influence from a source spreads into its future light cone), rather than depending on future conditions. This is the standard causal structure used in physically meaningful propagation. It is also an immediate way to build an arrow into the field definition without putting an arrow into the spacetime metric.

The payoff is visual: once you define propagation, light cones stop being mere boundaries. They become visible fronts—ridges that move forward in time. Behind those fronts can be wakes, decays, interference-like overlaps, and regions of elevated influence that look like a meteorological map of causality: waves, storms, and calm zones, all strictly constrained by Minkowski light cones.

6.1 From possibility to propagation: kernels, fronts, and wakes

To move from possibility to propagation, we need two ingredients: sources and a kernel. A source is a set of events that “emit” something—energy, information, disturbance, a signal, a causal impulse. A kernel describes how that emission affects other events. In the simplest retarded picture, the influence at event p depends only on sources in the past of p , with causal support restricted to the past light cone.

A general template is:

$$\Phi(p) = \int K(p, s) \rho(s) ds$$

Here $\rho(s)$ is the source density defined over spacetime (or over some subset such as a worldline), and $K(p, s)$ is a propagation kernel that encodes the rule “how much does a source at event s contribute to event p .” The essential Minkowski constraint is:

$$K(p, s) = 0 \text{ if } s \text{ is not in PastCone}(p)$$

That single condition enforces relativity’s causal structure. Everything else in the kernel is a modeling choice: attenuation with distance, frequency content, dispersion, scattering, noise, memory loss, or a simple uniform weight.

In many toy models, the most dramatic structure comes from the behavior near the light cone itself. The retarded front is the locus of events that are just becoming reachable from a source at the speed limit. In 1+1, this is a pair of lines $x = x_s \pm (t - t_s)$. In higher dimensions it becomes a spherical wavefront. When the kernel gives significant weight near this boundary, the diagram naturally produces sharp ridges—fronts that sweep forward like a causal shock.

Behind the front is a wake: the region already reached by the influence. The kernel can make the wake slowly decay, remain constant, or disperse. In causal-topography terms, the front is the “edge of new influence,” while the wake is the “region of influence persistence.” This distinction is visually intuitive and captures a real physical difference: wavefront arrival is often a detectable event, while the post-arrival field can be steady, oscillatory, or fading.

One can choose kernels that are intentionally simple for pedagogy. A step-kernel produces an influence plateau inside the future cone. A decaying kernel produces a smooth hill behind the wavefront. A peaked kernel produces sharp ridges along the cone. Each variant yields a different visual grammar, but all preserve the key Minkowski property: no influence outside the cone.

6.2 Visualizing sources: ridges on future cones rather than abstract boundaries

Standard Minkowski diagrams draw light cones as abstract boundaries associated with a chosen event. Model B turns those boundaries into dynamic features that appear because something is actually propagating.

Consider a single impulse source at event $s = (t_s, x_s)$. In Model B, you define a kernel K such that influence is strongest near the earliest arrival: the lightlike boundary of the future cone. Then Φ becomes elevated along the two lightlike rays emanating from s . In a 3D surface plot, these rays become ridges rising above the Minkowski plane, tracing out the future-directed “arrival lines.”

If instead you choose a kernel that fills the interior of the future cone, the ridge becomes less sharp and the interior becomes a raised wedge. That wedge expresses a simple message: after the wavefront passes, the region is causally downstream and remains influenced. If you then add attenuation, the wedge becomes a hill that slopes downward as you move away from the source, or downward as you move forward in time if the influence dissipates.

The key diagrammatic improvement is psychological: the light cone is no longer merely a restriction. It becomes a visible moving frontier separating “not yet affected” from “already affected.” In causal discussions, this is often what one wishes to communicate but cannot with a static 2D diagram. Model B supplies that missing dynamism without animating anything: the dynamism is encoded as a static 3D shape.

For multiple sources, the picture becomes even more interpretable. Each source contributes its own future-cone ridges and wakes. The surface height at any event becomes the accumulated influence. The Minkowski plane remains the stage, but the influence landscape shows where the stage is “lit up” by actual causal transmissions.

This also connects naturally to arrow-of-time intuition. A retarded model privileges future-directed influence by construction: effects depend on causes in the past cone. The advanced-time analog would flip this dependence and generate influence arriving from the future cone instead. Minkowski geometry allows both in principle as symmetric constructions, but the physical world overwhelmingly uses retarded behavior in macroscopic contexts. So Model B can display the arrow not by modifying spacetime, but by choosing the physically relevant causal Green’s-function orientation.

6.3 Superposition of influences and the emergence of causal “weather maps”

The most visually interesting part of Model B appears when you allow superposition. If Φ is defined as an integral or sum over sources, then multiple causal streams can overlap. The landscape becomes a map of causal intensity across spacetime.

In this setting, “causal weather” is a useful metaphor. Sources are storms or pressure systems. Wavefronts are moving fronts. Wakes are lingering humid zones. Quiet regions are deserts. The diagram lets you see causal structure as a spatially and temporally textured field rather than a binary constraint.

Several types of structure emerge naturally.

First, constructive overlap. Where multiple sources have future cones that intersect, Φ rises. This creates ridges or plateaus in the overlap region. Even without oscillatory interference, simple additive overlap produces “hotspots” of causal density.

Second, shielding and sparsity. If sources are localized and attenuation is strong, there will be regions inside the future cones that are only weakly influenced. These appear as low basins. This is important because it breaks the naive reading of a light cone interior as uniformly “reachable.” Reachable does not mean strongly influenced.

Third, moving observer effects. If you evaluate Φ along a particular worldline (an observer moving through spacetime), the observer experiences time-varying influence as fronts pass and overlaps change. The 3D surface makes this obvious: a worldline through the terrain is like a hike through weather, with sudden rises at fronts and slow declines in wakes. This creates a tight link between diagram and lived experience.

Fourth, causal clustering. If sources are distributed along a world-tube (representing, say, a lab emitting signals continuously), then instead of a single cone you get a layered set of cones whose superposition forms broad elevated regions. This looks like a causal “cloud” that trails forward in time. Such shapes begin to resemble the “accessible archive” domes of Model A, but with a different meaning: here the dome is produced by active emission rather than passive overlap.

Once superposition is allowed, you can ask new questions that standard Minkowski diagrams do not invite. Where in spacetime is influence concentrated? How does influence migrate? Where are the boundaries between regions dominated by different sources? How does attenuation reshape causal reach? These are physically meaningful questions in contexts ranging from signal propagation to radiation transport to social influence networks, provided the model is interpreted at the correct level of abstraction.

6.4 How this differs from field theory diagrams and why it remains Minkowski-based

At this point, one might object: isn’t this just drawing a field on spacetime, something that field theory already does? Why call it a “Minkowski-based diagram” rather than a generic spacetime field plot?

The difference is emphasis and grammar.

Field theory diagrams, in practice, usually assume the spacetime structure as background and focus on the dynamics of fields: equations of motion, propagators, mode expansions, gauge choices, and so on. The causal structure is present but often implicit. Even when one draws retarded Green’s functions, the goal is typically analytic: compute responses, prove support

properties, derive radiative behavior. The Minkowski diagram tradition, by contrast, is pedagogical and geometric: the goal is to make causal structure visually immediate. Light cones are drawn not as an afterthought but as the main organizing object.

Model B stays Minkowski-based because the light cone remains the central scaffold of interpretation. The retarded kernel is not introduced to solve a field equation in full generality, but to convert the Minkowski light cone from a boundary of possibility into a front of propagation. The diagram is still read in the Minkowski way: events, worldlines, cones, causal order. Φ is the overlay that adds intensity to that scaffold. It is a causal-structure-first visualization, not a dynamics-first calculation.

There is also a second distinction. Model B does not require the field you visualize to be a fundamental physical field. It can be an effective influence field: message rate, detection probability, record imprint strength, reliability-weighted signal flux, or any other operational quantity that propagates causally. This flexibility is precisely what makes the construction useful for discussing “information flow” and “records” without requiring the reader to commit to a specific microphysical ontology. You are not claiming that Φ is the electromagnetic field or a scalar quantum field. You are using the retarded causal rule as a template to define a scalar topography that represents influence in a broad, physically grounded sense.

Finally, Model B clarifies the arrow-of-time theme while respecting relativity. Minkowski geometry itself is symmetric. But the choice of retarded propagation is a physically motivated asymmetry: it encodes the fact that in ordinary macroscopic contexts, causes precede effects in the sense of influence traveling from past to future light cones. One could define an advanced model too, purely mathematically, but the contrast would only reinforce the paper’s core message: the arrow is not in the metric; it is in the additional physical structure—here, in the orientation of the propagation kernel and in the selection of boundary conditions.

In summary, Model B extends the Minkowski diagram from a static causal skeleton into a causal dynamics map by defining Φ as a retarded influence field. It turns cones into fronts, boundaries into ridges, and mere reachability into graded intensity. It creates a visual language in which causality behaves like weather: structured, overlapping, locally intense, globally constrained. And it does so without rewriting relativity, because the only absolute constraint it assumes is the one Minkowski geometry already supplies: causal support inside the light cone.

7. Model C: Entropy Production and Irreversibility Landscapes

Model A made the interior of the light cone non-uniform by asking how much of a record-bearing domain lies in an event’s causal past. Model B made it non-uniform by modeling actual influence propagation from sources and drawing retarded fronts and wakes. Model C goes to

the heart of what most people mean by “the arrow of time”: irreversibility. If the past and future cones are symmetric in Minkowski geometry, then the arrow cannot be a property of the cone. It must be a property of what happens inside the cone—specifically, a property of processes that convert low-entropy macrostates into higher-entropy ones and, in doing so, create robust records and destroy the fine-tuned correlations needed to reverse them.

The premise of Model C is that irreversibility is not a vague philosophical add-on. It is a physically localized phenomenon. Certain regions of spacetime are “arrow engines”: they are where entropy is produced, where microscopic distinctions are amplified into macroscopic outcomes, where energy is dispersed into many degrees of freedom, and where records become stable. Other regions are comparatively reversible: near-equilibrium, weakly dissipative, low interaction, low record formation. If we want a diagram that lets a reader see why the arrow is experienced as directional, then plotting an “irreversibility landscape” over Minkowski spacetime is a direct and intelligible move.

In this model, $\Phi(ct, x)$ is chosen to represent either a coarse-grained entropy associated with a subsystem at an event, or more sharply a local entropy production rate. The resulting 3D surface is no longer merely about causal reach or causal flux; it is about where the universe is actively converting causal possibility into irreversible history.

7.1 Φ as coarse-grained entropy S or local production rate σ

There are two natural choices for Φ in an entropy-based visualization.

The first is coarse-grained entropy S . Here we imagine that at each event (ct, x) we can associate a macrostate description of some subsystem—perhaps matter in a local region, perhaps an apparatus plus environment—and assign to it an entropy $S(ct, x)$ defined by a chosen coarse-graining. Then $\Phi = S$, up to scaling. Regions of spacetime where the subsystem is in a high-entropy macrostate rise higher in the landscape; low-entropy regions lie low.

This choice has pedagogical clarity: the height corresponds to “how many microstates are consistent with what we observe macroscopically.” However, it also has a conceptual drawback: entropy as a state variable can be high even in equilibrium where nothing is actively happening. A high S region is not necessarily a region where the arrow is being generated moment by moment; it may simply be a region that already sits at a typical macrostate.

The second choice is often more aligned with “arrow strength”: the local entropy production rate $\sigma(ct, x)$. This is a measure of how much entropy is being produced per unit time (or per unit proper time, depending on how one formalizes it) due to irreversible processes such as dissipation, friction, diffusion, chemical reactions, or irreversible amplification. In this reading, $\Phi = \sigma$. Now the height represents not a passive state but an active process: where the arrow is being manufactured.

This distinction is useful. S tells you “how far along” a region is in thermodynamic typicality; σ tells you “where the change is happening.” In many situations the most dramatic arrow-content is in σ rather than S . A chemical reaction zone can have large σ even if the final state is not dramatically higher in S than its surroundings; what matters is the irreversible conversion.

One can also combine them: $\Phi = \alpha S + \beta \sigma$, or $\Phi = \sigma$ with S used as a background. But for conceptual discipline, it is helpful to separate them. If the goal is to visualize irreversibility as a localized engine of directionality, σ is often the cleaner choice.

7.2 Entropy cliffs, ramps, and localized bumps as macro-amplification events

Once Φ is tied to entropy or entropy production, the geometry of the lifted diagram becomes a way to draw irreversibility as structure rather than as rhetoric. Several characteristic shapes become meaningful.

An entropy cliff is a sharp jump in Φ across a region of spacetime. In a toy model, this can represent a sudden macrostate transition: a metastable system snapping to equilibrium, a phase transition front, a detonation wave, a violent mixing event, or any process where an initially constrained configuration releases into many degrees of freedom rapidly. In the diagram, the cliff marks an irreversible boundary: “before” and “after” are not merely ordered in time; they are separated by an abrupt loss of micro-correlation.

An entropy ramp is a gradual slope. This represents slower, distributed irreversibility: diffusion of a gradient, thermal relaxation, slow mixing, gentle dissipation. The ramp’s slope is the local production: the steeper the ramp, the higher σ . The ramp also produces a directional visual cue: it literally slopes one way. But the key is that the slope is not a coordinate artifact; it is a plotted physical field.

A localized bump is a compact peak in Φ , often the most intuitive depiction of “a moment of irreversible amplification.” This is the form used in the earlier figure as an illustrative addition: a Gaussian-like bump centered at some (t_0, x_0) . It can represent an event where a microscopic difference is amplified into a macroscopic record. The bump is not “the event” in the Minkowski sense; it is the irreversibility intensity associated with the event, spread over a region because real systems have finite extent and finite response time.

Why are bumps so pedagogically useful? Because they correspond to something readers already understand: there are places in a causal history where something becomes irreversible. A decision point, a measurement, a macroscopic mark, a catastrophic mixing, a fracture. The bump says: here is where the universe paid the thermodynamic price that made a particular history robust. Outside the bump, events might still be causally linked, but history is “soft”; inside the bump, history becomes “hard,” copied, redundant, difficult to undo.

This is where Model C naturally complements Model A. Model A tells you how much causal archive is available. Model C tells you where archive is forged. A large past cone does not automatically produce robust records; it merely makes influence possible. Robust records require macro-amplification, and macro-amplification is what entropy production supplies.

7.3 The measurement analogy: microscopic symmetry, macroscopic irreversibility

One of the most illuminating ways to talk about irreversibility without drowning in formalism is the measurement analogy. The word “measurement” here should be understood broadly, not only as a quantum concept but as a general physical pattern: a small difference becomes a large, stable, redundant mark.

At the microscopic level, many interactions are reversible in principle. The underlying dynamics preserve information in a detailed sense. If you knew every microvariable and could reverse all momenta with perfect precision, you could in principle run the dynamics backward.

But macroscopic measurement-like processes are not simply micro-interactions. They are micro-interactions embedded in environments with enormous numbers of degrees of freedom. When a microscopic difference is amplified into a macroscopic record, the record is not stored in one variable; it is copied into many variables: photons scattered into space, phonons and heat in the apparatus, microscopic rearrangements in materials, neural patterns, digital bits, chemical changes. The environment becomes an echo chamber of the outcome. Reversal would require not only reversing the primary system but collecting and re-phasing the entire environmental spread. That is not forbidden by Minkowski geometry; it is suppressed by combinatorics.

This is the heart of irreversibility: microscopic symmetry does not survive macroscopic embedding. The equations can be reversible while typical macroscopic outcomes are not. The arrow is thus less a property of time itself and more a property of how information is allowed to disperse and how correlations become redundant.

In the Φ diagram, measurement-like events correspond to localized regions of high σ : arrow engines. They are the places where the universe turns an option among many microhistories into a single macroscopic branch that will now be remembered. The “branching” language can be used cautiously as a metaphor even without committing to any particular interpretation of quantum mechanics. The point is operational: after amplification, the world contains multiple correlated copies of a fact, and that correlation is thermodynamically stabilized. That is what makes “the past” feel fixed.

So the measurement analogy becomes a bridge: it connects the abstract concept of entropy production to the concrete concept of record creation. It also clarifies why an arrow-field Φ is not merely “entropy increasing” drawn as a chart. It is a map of where irreversible stabilization

happens. The arrow is not everywhere equally. It is concentrated at amplification points and dissipative pathways.

7.4 Why this can visualize an arrow without claiming time itself is generated

A common conceptual hazard in arrow-of-time discussions is to slip from “the arrow emerges” to “time is generated.” Those are different claims. The arrow is about directionality in macroscopic evolution and record structure. The existence of time as a coordinate in relativity is a separate structural feature of spacetime description. Model C is not attempting to derive time or to claim that time is an emergent illusion. It is attempting to show that within an already-given Minkowski framework, the arrow can be represented as a field-like asymmetry that comes from physical processes.

In this sense, Model C is conservative. It respects Minkowski geometry fully. Past and future cones remain symmetric in the base plane. The field Φ breaks that symmetry not by changing the geometry but by representing how the world actually populates the geometry under particular boundary conditions. If you reverse the time coordinate while keeping Φ fixed, you are no longer describing the same physical world-history; you are describing a different world-history with different boundary conditions. The arrow is not a gauge choice. It is a property of the state of the world and its record-bearing evolution.

This is exactly why the diagrammatic approach is valuable. It forces the reader to keep two truths simultaneously in view.

First, the causal skeleton is symmetric: Minkowski spacetime does not prefer a direction of time.

Second, the arrow-content is asymmetric: entropy production and record stabilization occur in a preferred temporal orientation because the world’s boundary conditions and typical dynamics select that orientation.

By lifting the Minkowski plane with an entropy-based Φ , we make the asymmetry visible without pretending it is intrinsic to the metric. The arrow appears as hills, slopes, and peaks—physical features of a field over spacetime, not metaphysical features of time itself.

The final conceptual advantage is that Model C can be combined cleanly with Models A and B. Accessibility (Model A) tells you how much of a record domain lies behind you causally. Influence (Model B) tells you how signals and disturbances actually propagate and accumulate. Entropy production (Model C) tells you where those propagations become irreversible and record-like. Together they motivate a single interpretive grammar: the light cone bounds possibility, influence fills possibility with flow, and entropy production locks some of that flow into history. The “third axis” Φ is the diagrammatic space in which these arrow-making mechanisms can be seen rather than merely asserted.

8. Symmetry, Invariance, and Interpretation Constraints

If we are going to add a third-axis field $\Phi(ct, x)$ to a Minkowski diagram and speak about it as causal accessibility, influence, or entropy-arrow strength, we owe the reader a discipline check: what is invariant, what is not, and what kinds of claims this diagram can and cannot support. Minkowski diagrams are beloved partly because they train the eye to respect Lorentz invariance. A careless Φ could undo that training by smuggling in frame-dependent quantities and presenting them as if they were geometric facts. The goal of this section is therefore not to defend Φ at all costs, but to constrain it. A good Φ is not one that tells a dramatic story; a good Φ is one that tells a story without confusing coordinates with physics, and without confusing observer conventions with invariant structure.

The deeper point is that “information” and “entropy” live in an awkward middle ground. They are physically grounded but description-dependent. They depend on what degrees of freedom are accessible, what coarse-graining is chosen, what counts as a record, and what counts as a system versus an environment. These choices are not arbitrary; they correspond to real operational constraints. But they do mean that objectivity in this domain is not the same as objectivity in the Minkowski metric. We must therefore develop a careful vocabulary: Minkowski geometry is the invariant causal skeleton; Φ is an added field that may be invariant in some constructions and deliberately frame-tied in others; and the diagram’s interpretation must keep those layers separate.

8.1 Can Φ be Lorentz-scalar, and when should it not be?

In principle, yes: Φ can be defined as a Lorentz scalar, but only if we define it in a way that does not rely on a preferred slicing of spacetime into “space at a time,” and does not rely on quantities that are themselves frame-specific. The simplest example is local scalar entropy production defined covariantly. In relativistic continuum thermodynamics, one can define an entropy current and its divergence; the divergence is a scalar that represents entropy production rate density. If you can define $\sigma(p)$ in this way for your chosen physical system, then $\Phi = \sigma$ is naturally Lorentz-scalar. In a Minkowski diagram, this would mean that while different inertial observers would draw different coordinate grids, they would agree on the value of Φ at a given event.

Similarly, certain accessibility measures can be made invariant if they are constructed geometrically using invariant spacetime volumes and invariant domains. For example, “the 4-volume of the intersection of a causal past with a specified spacetime region” is coordinate-independent, provided the region itself is defined in a coordinate-independent way. In 1+1 diagrams we casually use area; in full spacetime this becomes 4-volume. If D is defined as a physically specified world-tube (e.g., the world-tube of a particular material object), then the

measure of $\text{PastCone}(p) \cap D$ can be an invariant geometric quantity. Taking $\log(1 + \text{volume})$ preserves that invariance.

However, this is also where the trap lies. Many natural-seeming definitions of Φ implicitly choose a preferred frame. If you define D as “the region $|x| \leq a$ at times $0 \leq t \leq T$,” you have used a specific coordinate time t and a specific spatial coordinate x . That D is not invariant under boosts; a boosted observer will not describe the same set of events by those inequalities. In that case, $A(p) = \text{measure}(\text{PastCone}(p) \cap D)$ is not a scalar; it is tied to the chosen frame.

The correct conclusion is not “therefore Φ is illegitimate.” The correct conclusion is “therefore be honest about what Φ is.” If Φ is meant to represent accessibility relative to a laboratory at rest in a particular frame, then frame dependence is not a bug; it is precisely what you want. Laboratories have rest frames. Clocks, detectors, and memory-bearing hardware define a preferred world-tube. That is physical. Lorentz invariance is not violated by acknowledging that a particular apparatus has a particular motion state.

So the rule is: Φ can be Lorentz-scalar when it is built from covariant local quantities or invariant geometric constructions; Φ should not be forced to be Lorentz-scalar when it is meant to represent an operational quantity defined relative to a particular observer or apparatus.

8.2 Frame dependence as a feature: observers, records, and operational meaning

Many of the most meaningful “information” quantities are operational: they describe what an observer can in fact access, store, transmit, or infer. Operational meaning is often inseparable from an observer’s worldline and rest frame.

Consider “record accessibility.” If D is “my laboratory’s world-tube,” then accessibility is literally observer-relative. Two observers in relative motion do not share the same lab. They do not share the same stored records. They do not share the same causal pathways through which information reaches their instruments. Expecting Φ to be invariant in that setting would be like expecting “what is in my filing cabinet” to be Lorentz-invariant. It is not. It is a property of a physical subsystem.

Similarly, influence propagation in Model B can be made covariant at the level of field equations, but the moment you translate it into a scalar “influence density” that is meant to represent detectability, signal-to-noise, or effective impact, you are likely to introduce frame-tied elements: detector response bandwidth, time integration windows, sampling rates, thresholds. Those are physical. A moving detector sees Doppler shifts and time dilation. If Φ is meant to represent “what the detector registers,” then it should transform accordingly.

Frame dependence becomes a feature when you interpret Φ as a map of lived causality rather than abstract causality. Minkowski geometry is the grammar of abstract causality. Lived causality

is mediated by matter that carries clocks, memories, and sensors. Those devices select frames and coarse-grainings. They are not optional decorations; they are the mechanism by which information becomes real in a world.

The diagram therefore offers a useful conceptual separation: causal structure is invariant; record structure is worldline-anchored. If the reader accepts that, then a frame-dependent Φ is not an embarrassment. It is a precise way of saying: “given this observer, this apparatus, and this record domain, here is the causal topography that matters to them.”

This also connects to the arrow theme. The arrow of time is not merely “time labeled upward.” It is the existence of stable records and irreversible macrostates. Those records exist in physical substrates that move along worldlines. There is no record “in general”; there are records for someone or something. So in many realistic uses, Φ is properly interpreted as an observer-coupled field.

8.3 Avoiding category errors: Minkowski geometry versus added physical structure

A category error is the primary failure mode of this entire paper. There are two forms to guard against.

The first is to treat Φ as if it were an extra spacetime dimension. It is not. It does not change the metric. It does not create new causal pathways. It is a plotted value that lives on events. The underlying causal relations are still determined by Minkowski separations. The light cone remains the boundary of influence; Φ does not expand or shrink it.

The second is to treat Φ as if it were a derived necessity of Minkowski geometry. It is not. Minkowski geometry permits many Φ fields, including the trivial one $\Phi = 0$ everywhere. The existence of a nontrivial Φ is a statement about the physical world’s content and history, not a statement about spacetime kinematics. In particular, if Φ is entropy production or record formation, it depends on matter, interactions, and boundary conditions.

This is why the paper insists on the language “skeleton plus overlay.” Minkowski is the skeleton. Φ is overlay. They are related, but not identical. Minkowski determines what causal intersections are possible; Φ chooses which of those possibilities are populated by record-bearing structure, influence, and irreversibility.

Keeping these categories separate yields a disciplined interpretive practice. When you see a ridge in Φ , you do not infer “causality is stronger there” in the geometric sense. You infer “given the chosen definition, record accessibility or influence intensity is higher there.” That is a statement about a field on spacetime, not about spacetime itself.

This separation also prevents the common rhetorical slide where one says “the arrow is built into spacetime.” The diagrammatic approach actually argues the opposite: by drawing the arrow

as Φ above spacetime, you visually demonstrate that the arrow is not the base geometry. It is something added.

8.4 What “objectivity” could mean for information and entropy in this diagram

Objectivity in Minkowski space is crisp: the invariant interval, causal order for timelike separations, and the light cone structure are objective in the strongest sense. Any inertial observer agrees on them.

Objectivity for entropy and information is subtler. It is not “anything goes,” but it is also not absolute in the same way as the metric. What is objective, and in what sense, depends on the level of description.

One kind of objectivity is operational reproducibility. If different observers build the same apparatus, calibrate it, and run the same protocol, they should obtain consistent Φ values in their own frames. Even if the numerical representation changes under boosts (because the apparatus response changes), the protocol-defined quantity remains well-defined. This is the kind of objectivity that most of experimental physics lives on: not coordinate-free metaphysics, but reproducible operational meaning.

A second kind of objectivity is covariant definability. If Φ is defined as a Lorentz-scalar derived from local covariant fields (for example, the divergence of an entropy current), then Φ has a stronger objectivity: it is the same number at the same event for all inertial observers. This is an ideal, but it is not always the right ideal when Φ is meant to represent observer-limited access to records.

A third kind of objectivity is stability under reasonable coarse-graining. Entropy is coarse-grained by nature. But if a proposed Φ changes wildly under small changes of coarse-graining that do not matter physically, it is not robust. Conversely, if Φ captures a stable feature—such as “this process is strongly dissipative” or “this region contains dense record-bearing structure”—then it has a pragmatic objectivity: it picks out a real feature of the physical situation, even though the exact numeric value depends on choices.

A fourth kind of objectivity is relational invariance. Some information statements are best understood relationally: “event p is downstream of record domain D ,” “records become redundant after region R ,” “this worldline intersects a high- σ region.” These relational statements can be invariant even when the detailed field values depend on frame. The diagram can therefore support objective claims at the level of structure—where are the ridges, where are the basins, where are the fronts—without insisting that every numerical height is observer-independent.

This layered view of objectivity is not an evasion; it is the correct epistemic posture for a diagram that mixes a coordinate-free causal skeleton with a description-dependent arrow field. Minkowski geometry gives hard objectivity. Φ gives context-bound objectivity: covariant when possible, operationally reproducible when not, and structurally robust when the definition is well-chosen.

The practical outcome is a set of interpretation constraints that can be stated simply. Never confuse Φ with a coordinate. Never confuse Φ with the metric. Be explicit about whether Φ is intended to be Lorentz-scalar or observer-coupled. Demand operational meaning and robustness under reasonable modeling changes. And treat objectivity for Φ as a layered notion—sometimes numerical, sometimes structural, sometimes relational—appropriate to fields that represent records, information, and irreversibility rather than pure geometry.

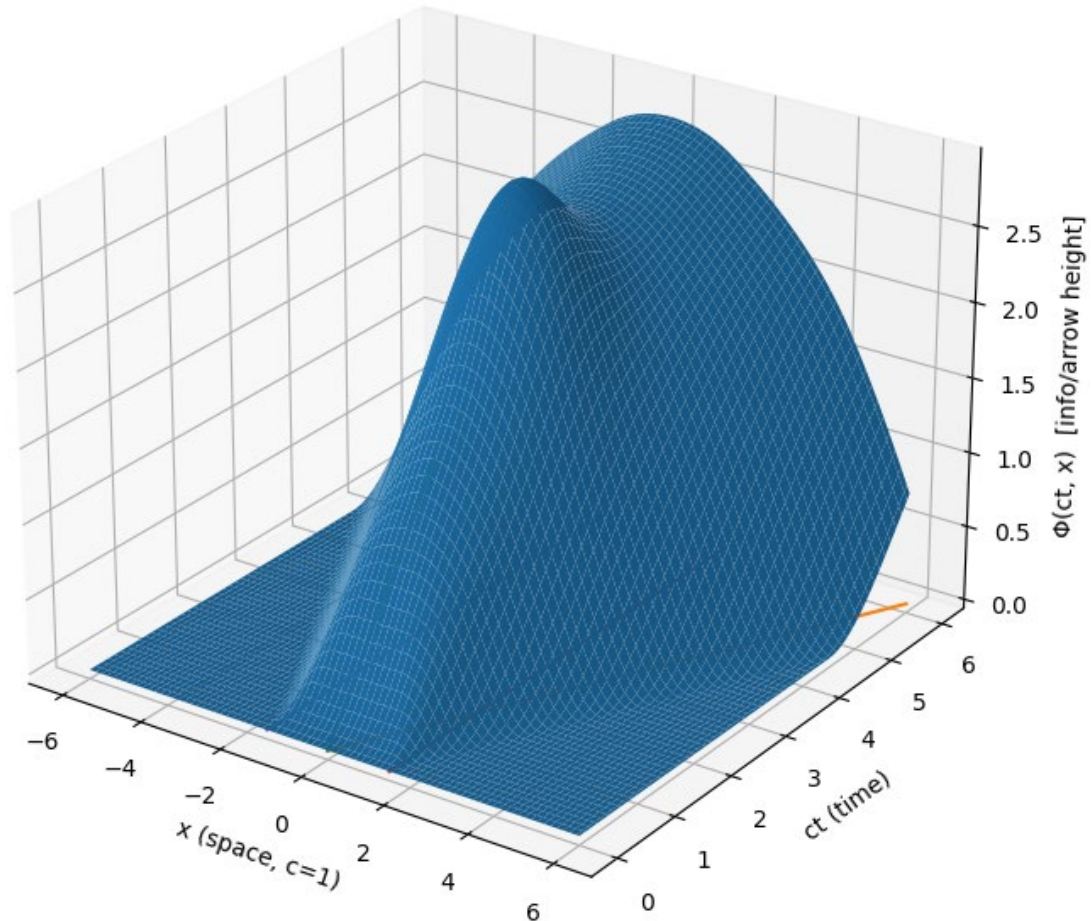
With these constraints in place, the lifted Minkowski diagram becomes intellectually honest: it preserves the invariants that relativity demands, while still giving the reader a way to see why real histories are asymmetric inside a symmetric causal skeleton.

9. Worked Example: The 3D Lifted Diagram and How to Read It

This section treats the lifted diagram as an instrument rather than as a metaphor. The goal is to show how a reader can “read” the figure the way one reads a contour map or a weather chart: not by staring at it until it feels philosophical, but by learning what each element means and how to extract concrete statements from the shape. The figure introduced earlier is a deliberately simple composite: it uses Model A as the baseline (causal accessibility as past-cone overlap with a record domain) and then adds a localized “entropy production” bump to demonstrate how irreversibility can appear as a localized intensification on top of the accessibility landscape. Nothing in the figure requires new physics; it is a visualization built from Minkowski causal geometry plus a physically motivated overlay field.

In this section, the important move is to connect diagram-reading to specific questions: Where is causal access to records high? Where does it saturate? Where does irreversibility concentrate? How would different observers or different record domains change the topography? Once the reader can answer those questions from the picture, the diagram has succeeded.

3D Minkowski-style diagram with a causal/information/entropy 'height' field
 $\Phi = \log(1 + \text{area}(\text{PastCone} \cap \text{world-tube})) + \lambda \cdot (\text{localized entropy-production bump})$



9.1 Diagram components: base plane, light cone lines, record world-tube boundaries

The diagram should be read in layers.

Layer one is the base plane: the familiar Minkowski diagram in 1+1 dimensions. One horizontal axis is x (space) and the vertical axis is ct (time) with units chosen so that $c = 1$ (or equivalently ct is used instead of t). Every point on this plane is an event. Every curve drawn on the plane (if present) would represent a worldline. The base plane is the causal skeleton: it is where you apply the basic Minkowski rules about timelike, spacelike, and lightlike separation.

Layer two is the light cone reference lines. In the figure, two straight lines on the floor plane emanate from the origin with slopes corresponding to $x = \pm t$. These are the lightlike boundaries of the cone. They are not “the only signals”; they are the boundary of maximum causal reach.

Inside that V-shaped region (in 1+1) lie events that can be influenced by the origin. Outside it lie events spacelike separated from the origin.

Layer three is the record-bearing domain boundaries. In this specific worked example, D is chosen as a world-tube centered near $x = 0$ with boundaries at $x = \pm a$. On the base plane, these appear as two vertical-in-time lines. They mark “where record-bearing structure lives” in the simplified model: a spatial strip that persists over time. Depending on interpretation, this tube could represent the laboratory, a sensor network, a memory-bearing environment, or any region where “records exist and signals are generated.” The key is that D is not everywhere; it is localized. That localization is what turns the interior of a light cone into a non-uniform accessibility landscape.

Layer four is the height axis itself: $\Phi(ct, x)$. This is the “not spatial” axis. Height does not represent another direction in space; it represents a scalar quantity assigned to each event. In the worked example, Φ is a sum of two terms: (i) a baseline accessibility term derived from past-cone overlap with D , and (ii) a localized bump representing enhanced irreversibility (an entropy-production-like event) centered at a particular spacetime location.

A reader should pause here and internalize the grammar: everything on the base plane is relativity; everything in height is added physical structure. The light cone lines and the tube boundaries are guides that help interpret why the surface rises where it rises.

9.2 Interpreting the surface: low- Φ deserts, high- Φ plateaus, entropy bumps

Once the components are understood, the surface becomes readable as terrain.

Low- Φ deserts are broad regions where the surface remains near the base plane. In Model A terms, these are events whose past light cones intersect little or none of the record domain D . This can happen because the event is too early in time (its past cone is shallow) or too far in space (its past cone does not reach the tube within the available time window). In the figure, you see low-lying regions at early ct and at large $|x|$. The interpretation is: even though Minkowski geometry permits causal past, the event has little causal contact with the designated record-bearing region.

High- Φ regions appear as hills or plateaus. These correspond to events whose past light cones intersect a large portion of D . Near the center of the tube and at later times, the overlap grows: as time advances, an event’s past cone opens up and captures more of the world-tube’s history. This produces a rising surface that can saturate into a plateau if D ’s time extent is finite. That plateau is a diagrammatic expression of “accessible archive saturation”: once the past cone already includes essentially all of the tube’s history, later times do not add much new accessible tube-volume.

In this worked figure, you also see a localized bump riding on top of the baseline hill. This bump represents a region where Φ is increased not because more of D is accessible, but because an additional term has been added to represent localized irreversibility. Think of it as “the arrow engine”: a dissipative, amplifying, record-forming event. It is intentionally drawn as a smooth Gaussian-like peak to indicate that real irreversibility is not a mathematical point but a finite region of space and time where macroscopic amplification occurs.

The key interpretive discipline is to separate these two meanings. The baseline hill is about how much causal past could have reached the event (relative to D). The bump is about where irreversibility is concentrated. Both elevate Φ , but they represent different physical stories. The diagram invites you to see them together: a region can be highly accessible yet not strongly irreversible, or strongly irreversible yet not globally accessible, depending on how the terms are chosen.

9.3 Diagnostic readings: where information is available, where arrow-strength concentrates

A useful diagram is one that helps you answer diagnostic questions. Here are the kinds of readings the worked example supports.

First diagnostic: “Where is information available, in the sense of accessible records?”

Look for high baseline Φ . In Model A, high baseline height means a large overlap between $\text{PastCone}(p)$ and D . This identifies events that sit downstream of a rich record domain. If D represents the lab or observer memory, this is literally “where the observer has a large accessible causal archive.” If D represents matter-rich environment, it is “where the event sits causally downstream of structured degrees of freedom.” In either case, you read availability as elevation above the deserts.

Second diagnostic: “Where does the arrow-strength concentrate?”

Look for sharp localized peaks or steep slopes if Φ includes an entropy production term σ . A localized bump is a signature of concentrated irreversibility: a region where microscopic distinctions become macroscopic records and where reversal becomes statistically suppressed. In the worked example, the bump tells you not only that “something happened there,” but that “something happened there that hardens history.” This is different from mere causal contact.

Third diagnostic: “Is the arrow evenly distributed or patchy?”

In ordinary Minkowski diagrams, the arrow is invisible, which encourages a mental picture that time’s direction is uniform. But in physical reality, irreversibility is often patchy: a quiet region near equilibrium has low σ , while a reaction zone has high σ . The lifted diagram makes patchiness visible. It teaches the reader to think of the arrow as a field-like phenomenon: it can be strong here and weak there, even inside the same causal cone structure.

Fourth diagnostic: “How do accessibility and irreversibility interact?”

A region of high accessibility can host a bump of irreversibility, and that bump can produce robust records that later become part of the accessible archive. The diagram can be read narratively in a disciplined way: the bump is where records are forged; the plateau is where those records are widely accessible later. In a more elaborate model, one could even define Φ so that entropy-production events in D increase record density in D , thereby feeding forward into Model A. Even without that feedback explicitly included, the layered terrain suggests a causal logic: irreversibility creates durable structure; durable structure increases future accessibility.

Fifth diagnostic: “Where are causal horizons in practice?”

The light cone gives absolute horizons, but the surface gives practical horizons. A region may be inside the cone yet still low- Φ because there are no records, no sources, or high noise. The diagram teaches a subtle point: Minkowski reachability is necessary for influence but not sufficient for meaningful information. Practical causal horizons can exist inside the geometric cone. The surface makes that visible as “dry basins inside the wedge.”

These diagnostics reveal why this is not just an artistic extension. It is a tool for reasoning: it encourages a more nuanced mental model of causality as “constraint plus content.”

9.4 Extensions: multiple domains, competing observers, and moving record-tubes

The worked example is intentionally simple: one world-tube, one baseline accessibility term, one entropy bump. But the framework generalizes naturally, and the generalizations are where the diagram becomes a genuine research and pedagogy tool.

Multiple domains: Instead of one record tube D , define several domains $D_1, D_2, \dots$ representing different laboratories, communication networks, or matter-rich regions. Then define Φ as a sum or a max of accessibility measures to each domain, possibly with different weights. Visually, this produces multiple hills whose overlaps form saddles and ridges. The terrain becomes a map of “which record domain dominates accessibility here.”

Competing observers: Different observers have different world-tubes and different coarse-grainings. You can draw separate Φ fields for each observer or define a vector of fields (Φ_A, Φ_B) and compare them. Even without plotting both surfaces at once, the concept clarifies an important point: “accessible past” is not universal in practice. It is conditional on the physical substrate of memory and detection. Two observers can inhabit the same Minkowski spacetime and yet live on different causal terrains.

Moving record-tubes: If the record domain is in motion, D is no longer a vertical strip; it becomes a slanted tube representing a moving apparatus. This automatically changes the overlap geometry and therefore the Φ landscape. This is not a bug; it is a direct way to visualize

how motion changes access to records and signal pathways. It also provides a clean bridge to relativity pedagogy: students can see that the causal skeleton stays invariant while the operational record domain changes its relationship to the cones.

Time-dependent or decaying domains: Real records degrade. Communication channels fail. Matter distributions change. You can represent this by weighting D as a function of time: dense early, sparse later, or vice versa. Then Φ becomes a map not only of causal reach but of record fragility. The plateau might disappear; the hill might erode; the deserts might expand.

Multiple irreversibility sources: Instead of one bump, include multiple bumps or extended ridges representing distributed dissipation. Then the arrow field becomes a landscape of many irreversible “hotspots.” This helps depict a world in which irreversibility is not a single event but a network of processes.

Hybrid Model A–B–C synthesis: Finally, you can define Φ as a weighted sum of accessibility, influence flux, and entropy production. The surface then has a layered meaning: it rises where the past is accessible, where influence is strong, and where irreversibility is active. Such a composite Φ is not a single physical observable; it is a deliberate visualization that integrates multiple arrow-relevant aspects of reality into one interpretable terrain.

The guiding principle remains the same across all extensions: the Minkowski base plane remains the invariant causal skeleton; Φ remains an overlay field whose definition is explicit and whose interpretation is disciplined. The reader’s reward is a diagram that can express, in a single image, a truth that is hard to hold mentally: the causal structure of spacetime is symmetric and invariant, but the history written inside that structure is asymmetrical, record-dependent, and physically textured.

10. Pedagogy and Research Utility: What This Adds Beyond Standard Diagrams

The lifted Minkowski diagram is not intended as a replacement for standard spacetime diagrams. The standard 2D Minkowski diagram is already close to optimal for what it was designed to do: make invariant causal structure visible. Its power comes from its austerity. But that same austerity leaves a conceptual gap that repeatedly surfaces in teaching and in informal reasoning: students and researchers can learn causal possibility while still remaining confused about causal texture—about why real histories have direction, why records accumulate, why some influences become durable and others evaporate into noise, and why “the future” is not simply the past reflected upward.

The proposal of a third, non-spatial axis $\Phi(ct, x)$ is best judged by whether it helps fill that gap without committing a category error. If it enables the reader to keep relativity’s causal skeleton

and thermodynamics' arrow-structure in the same mental frame—without collapsing one into the other—then it has pedagogical value. If it also suggests concrete questions, diagnostics, and experimental or modeling targets—where to expect record robustness, where to expect fragility, where irreversible amplification should cluster—then it has research utility. This section argues for both, while also naming the limits and failure modes that must be owned explicitly if Φ is to remain intellectually honest rather than merely evocative.

10.1 Teaching causal structure plus arrow-structure in one unified picture

A recurring teaching failure mode in relativity is that learners acquire the “light cone rule” but not the lived meaning of it. They can say which events are timelike separated, but they still imagine the interior of the cone as a uniform region where “the same kind of causality happens everywhere.” In real systems, causality is highly uneven. It is mediated by matter, by channels, by attenuation, by noise, by memory substrates, and by irreversible amplification. A bare Minkowski diagram does not show any of that. It is like drawing coastlines without elevation: you understand boundaries but not terrain.

The lifted diagram provides a unified picture precisely by keeping the original base plane unchanged. The student still sees the light cone, still learns the timelike/lightlike/spacelike partition, still understands that the cone is invariant and defines causal possibility. But now, in the same view, the student also sees that inside the cone there is a landscape: regions that are “record-rich” or “influence-intense” or “irreversibility-active.” The diagram therefore teaches two different but compatible ideas at once:

- The cone sets what is possible.
- The field Φ suggests what is typical, accessible, detectable, or irreversible.

This is pedagogically valuable because many confusions about time arise from mixing these two ideas. Students often treat what is possible as what is likely, or treat geometric symmetry as physical symmetry of lived histories. In a single picture, you can show: yes, the future and past cones are symmetric as boundaries; no, the landscape of records and irreversibility need not be symmetric. The difference becomes visible.

The diagram also supports a more mature intuition: that “the past” in practice is not the full past cone but the recorded past—what has left traces accessible to an observer or to a domain. By tying Φ to record domains and entropy production, you teach the distinction between causal past and epistemic past without turning it into a purely philosophical aside.

Finally, the picture can make the arrow of time feel less like a metaphysical primitive and more like a physical field: strong in some regions, weak in others. That is closer to how real physics

behaves. Dissipation and amplification are localized. Equilibrium regions are arrow-quiet. The lifted diagram makes this locality visually obvious.

10.2 Bridging relativity and thermodynamics without collapsing one into the other

There is a longstanding conceptual tension between relativity and thermodynamics that appears in informal discourse more than in formal theory. Relativity treats time as a coordinate woven into spacetime geometry; thermodynamics treats time as directionally loaded because of entropy and irreversibility. Many readers—especially those encountering relativity visually for the first time—attempt to resolve the tension by forcing one framework to absorb the other. They either say, “thermodynamics must be encoded in spacetime geometry,” or they say, “if spacetime is symmetric, the arrow must be illusory.”

The lifted diagram offers a third approach: layer them. Minkowski gives the invariant causal skeleton; thermodynamic arrow-content is an added physical structure living on that skeleton. This layering accomplishes a bridge without a collapse. It respects the difference between:

- Kinematics: what transformations preserve causal structure (relativity).
- State and dynamics: what macrostates are typical and what processes are irreversible (thermodynamics).

In practical terms, the diagram helps the reader see why the arrow of time is not “in” Minkowski geometry, but can still be plotted “on” Minkowski spacetime. That is exactly the right relationship if the arrow arises from boundary conditions and dissipative physics rather than from the metric.

This bridging is also useful for preventing a common conceptual mistake: confusing causal order with thermodynamic direction. Relativity gives you invariant causal order only for timelike relations; thermodynamics gives you a direction of macroscopic evolution. They are related—records and causes line up in the world we live in—but they are not identical concepts. A layered diagram makes it easier to teach that distinction cleanly.

10.3 Visualization as hypothesis generator: where to expect records, fragility, or amplification

Beyond pedagogy, the lifted diagram can act as a generator of hypotheses, especially in complex systems where causal accessibility, influence propagation, and irreversibility interact.

If Φ is defined by Model A (past-cone overlap with record domains), then the diagram highlights where information could be available. That suggests hypotheses about epistemic asymmetry: where observers should have rich reconstructable pasts versus where they should be causally impoverished. In social systems, this resembles the difference between regions embedded in communication networks versus isolated regions. In physical systems, it resembles the

difference between matter-rich, interaction-rich environments versus sparse, weakly interacting regions.

If Φ is defined by Model B (retarded propagation), then ridges and wakes highlight where influence should be strongest. That can suggest hypotheses about detectability and effective causal impact: where signals should be measurable, where they should be drowned by attenuation, where overlapping sources create hotspots, and where “quiet basins” persist even inside geometric reachability. This is particularly helpful in any domain where causal reach is not enough—where strength matters.

If Φ is defined by Model C (entropy production), then peaks and steep slopes highlight where irreversibility concentrates. That suggests hypotheses about record formation and fragility: where macroscopic “facts” should become stable, where memory should be robust, where irreversible events should seed long-lived constraints on future evolution, and conversely where the world should remain thermodynamically quiet and hence less record-generating.

The real research utility comes when these layers are combined. A region might have high accessibility (Model A) but low entropy production (Model C), implying a rich archive but little current amplification. Another region might have low accessibility but high entropy production, implying intense local irreversibility but poor dissemination into the broader record network—an arrow engine that does not become widely known. Those contrasts are not merely diagrammatic; they point toward concrete questions about how information and irreversibility couple.

In this way, the diagram behaves like a scientific sketch: not a proof, but a map that invites targeted questions. Where are the ridges? Why do they form there? What parameters control their sharpness? What happens if you move the record tube? What happens if the entropy bump becomes extended? What happens if two record domains compete? Each question corresponds to an interpretable change in the field definition, and therefore to a family of models one could explore analytically or computationally.

10.4 Limits, failure modes, and what would count as a better definition of Φ

The same flexibility that gives Φ its usefulness also creates its biggest vulnerability: Φ can become a Rorschach test. If it is defined too loosely, the diagram risks becoming a persuasive picture with no stable meaning. Therefore, the approach must come with explicit limits and criteria for improvement.

A first limit is non-uniqueness. There is no single canonical Φ . Accessibility, influence, and entropy production are related but not identical, and each can be quantified in many ways. This is not fatal, but it means the paper must present Φ as a family of definitions rather than as a universal new variable.

A second limit is coarse-graining dependence. Any entropy-like Φ depends on what variables are ignored and what subsystem is chosen. If the diagram is used carelessly, one can confuse coarse-graining artifacts with physical structure. A better Φ is one whose qualitative features (ridges, plateaus, hotspots) are robust under reasonable coarse-graining changes.

A third limit is frame dependence. Some Φ choices can be Lorentz scalars; others are naturally tied to an observer or apparatus. The failure mode is not frame dependence itself but implicit frame dependence. A better Φ definition either is covariant by construction or explicitly states the operational frame and why that is physically appropriate.

A fourth limit is causal cheating. Φ must not quietly imply superluminal influence unless that is the explicit topic. Influence kernels must have support inside the light cone. Accessibility measures must be built from causal past intersections. Entropy production terms should not create interpretive suggestions of “information arriving from outside the cone.” A better Φ respects Minkowski causality at the definitional level.

A fifth limit is interpretive overload. If Φ is defined as a weighted sum of many things, the diagram can become unreadable: a single peak might be caused by accessibility, influence overlap, or entropy production, and the reader cannot tell which. A better Φ supports diagnostic decomposition: either plot components separately or define Φ in a way that preserves interpretive clarity (for example, baseline accessibility plus a clearly labeled localized entropy term).

So what would count as “better” in this framework? Several criteria are reasonable.

- Covariance when possible: If Φ is meant to represent a physical production rate, prefer covariant definitions (e.g., a scalar entropy production density) that do not depend on arbitrary slicing.
- Operational clarity when not: If Φ is observer-relative, define it through a protocol—what is measured, with what apparatus, along what worldline.
- Robustness: The landscape should not change qualitatively under small, reasonable modeling variations.
- Causal fidelity: Φ should respect light-cone support and never confuse “inside the cone” with “actually influential” without stating the kernel or weighting.
- Pedagogical parsimony: The definition should be as simple as possible while still illuminating the intended concept.

Finally, the most important failure mode to avoid is the metaphysical overreach: reading Φ as if it were “time itself” or “a new dimension.” The diagram becomes better—not worse—when it

remains modest: Minkowski gives causal possibility; Φ visualizes arrow-content as a field living on spacetime. If the reader comes away with that layered picture, the diagram has done its job.

In short, the lifted Minkowski diagram adds something standard diagrams do not: a way to keep the causal skeleton and the arrow-structure simultaneously visible, while staying disciplined about what belongs to geometry and what belongs to the contingent physical world that occupies that geometry. It is a pedagogical bridge and a research sketchpad, provided Φ is defined with explicit constraints and interpreted with humility.

11. Open Problems and Future Directions

The lifted Minkowski diagram is intentionally a “first instrument.” It takes a familiar causal skeleton and adds a field Φ that makes arrow-relevant structure visible. But the moment we do this, we expose a cluster of deeper questions—some technical, some conceptual, some methodological. In this section, “open problems” does not mean “we don’t know physics.” It means: once you decide to visualize causal texture, you must decide how to define it, how to generalize it beyond 1+1, how to keep it honest under coordinate changes, and how to make it useful across domains without letting it dissolve into aesthetic metaphor. The promise is real: a unified picture of causal constraint plus arrow-structure. The work ahead is to make that promise portable, falsifiable in operational contexts, and modular enough to support multiple interpretations of Φ without conceptual drift.

11.1 Generalizing to 3+1: from surfaces to volumes and isosurfaces of Φ

The first nontrivial step is leaving the pedagogical comfort of 1+1. In 1+1, the diagram is a plane and Φ is a height field. In 3+1, the natural domain is four-dimensional, so there is no literal “height” that can be plotted without choosing a projection. This is not merely a visualization inconvenience; it forces us to confront what we actually mean by Φ .

One practical pathway is to treat Φ as a field on 3+1 spacetime and then visualize it through slices and isosurfaces:

- Choose a foliation into time slices (a family of spatial hypersurfaces). On each slice, you can plot Φ as a 3D volume (a scalar field over space) and display it using volume rendering or by drawing isosurfaces (surfaces where Φ equals a constant value).
- Alternatively, fix a spatial plane or a set of worldlines and plot Φ along them, producing a family of 2D “lifted” surfaces analogous to the 1+1 picture.

- Or choose causal objects—like causal diamonds associated with observers or events—and summarize Φ over those objects, producing scalar summaries that can be plotted in lower dimensions.

The open problem is not “can we render it” but “which rendering corresponds to the intended meaning.” In 1+1, it is easy to think that height is intrinsic. In 3+1, you must decide whether you want Φ to be read locally (value at an event), regionally (integral over a neighborhood), or relationally (value associated with a causal set like a past cone intersection). Each choice has consequences for both invariance and interpretation.

A major technical issue is the choice of measure in Model A when generalized to 3+1. The intersection $\text{PastCone}(p) \cap D$ becomes a four-dimensional region. Its 4-volume can be invariant if D is defined geometrically (a world-tube of a physical object), but in many practical uses D will be defined by an apparatus and a lab frame. In that case, the “volume of overlap” becomes observer-tied. That may be appropriate, but it must be explicit. In 3+1, this issue is sharper because “volume” and “time window” choices are more numerous and less visually obvious than in 1+1.

A second major issue is the shape of influence fronts in Model B. In 1+1, wavefronts are lines; in 3+1, they are expanding spheres (or more generally null hypersurfaces). A “ridge on the light cone” becomes an isosurface on a null shell. When multiple sources superpose, the geometry of overlaps becomes much richer: intersecting null shells create curved lines and surfaces of elevated Φ . In practical terms, the diagram becomes more like a 4D meteorology problem: causal “storms” propagate on null surfaces and fill timelike interiors. A toolkit would need conventions for showing these structures without losing the causal intuition that Minkowski diagrams provide.

A third issue is pedagogy: in 3+1 you risk losing the immediate clarity of light cones as drawn boundaries. If the visualization becomes too volumetric too quickly, it can obscure the simple causal partition it was meant to preserve. A promising compromise is to keep the 1+1 lifted diagram as the “explanatory interface,” but generate it as a slice through a true 3+1 model. Then the 1+1 picture is not the physics; it is the window into the physics.

Finally, there is an open design question: what are the right “primitive visual objects” in 3+1 causal topography? In 1+1, the primitives are cones, lines, and surfaces. In 3+1, the primitives might become:

- isosurfaces of Φ (arrow-strength shells)
- null hypersurface fronts (arrival boundaries)
- observer world-tubes (record domains)

- causal diamonds (bounded causal regions)
- horizon-like boundaries (regions of causal exclusion)

Choosing a small set of primitives—and teaching readers how to interpret them—may be the single most important step in making the framework scalable.

11.2 Curved spacetime: replacing cones with causal diamonds and horizons

Moving from Minkowski spacetime to curved spacetime changes the geometry of causality while preserving its logic. Light cones still exist locally, but globally the causal structure can be distorted by curvature, gravitational time dilation, horizons, and global topology. This creates both a challenge and an opportunity.

The challenge is that the clean “45-degree light cone” is no longer globally meaningful. In curved spacetime, the boundary of causal influence is determined by null geodesics, and their shapes depend on the metric. So the first step in generalization is to replace “cones” drawn as straight lines with causality determined by the spacetime’s causal structure: the set $J^-(p)$ (causal past) and $J^+(p)$ (causal future). In practice, for visualization, you would compute or approximate causal diamonds: intersections of a future region of one event with the past region of another. In many contexts, causal diamonds are more useful objects than cones because they are bounded and thus easier to measure and compare.

The opportunity is that curved spacetime naturally introduces “arrow-like” global structures that can interact with Φ in meaningful ways:

- Horizons restrict causal accessibility. Even if entropy production is high somewhere, horizons can prevent the resulting records or influences from reaching certain observers. In a lifted diagram, this would appear as a high- Φ region that does not translate into high accessibility for some worldlines.
- Gravitational time dilation changes the operational meaning of rates like entropy production per unit coordinate time versus per unit proper time. A covariant σ defined per unit spacetime volume can remain a scalar, but an operational “per tick of my clock” rate will differ along different worldlines. This makes the distinction between Lorentz/covariant scalar Φ and observer-defined Φ more than a philosophical aside; it becomes unavoidable.
- Curvature can focus or defocus geodesics, effectively changing the “density” of causal channels. Even without changing any microphysics, the geometry can make influence pathways converge, potentially increasing effective causal intensity in certain regions. In Model B language, the kernel’s spatial spreading changes because null geodesics spread

differently. That could generate causal “lensing” features in Φ : ridges sharpened by focusing, or broadened by defocusing.

An open problem here is how to keep the diagram’s interpretive discipline: we must not confuse geometric constraints (curvature and horizons) with thermodynamic arrow production. The best way to handle this is to treat curvature as part of the base skeleton (the causal structure) and treat Φ as an overlay that can be defined relative to that skeleton. Then a horizon is not a high ridge in Φ ; it is a structural boundary of reachability that changes how Φ is read by different observers.

There is also a methodological question: in curved spacetime, “record domains” D may themselves be best defined as world-tubes of matter distributions that gravitate. That means D is coupled to the geometry. In Minkowski space, we could choose D freely as a pedagogical object. In general relativity, matter and geometry co-determine causal structure, so the domain that produces records also shapes the causal pathways by which those records propagate. A mature causal-topography framework would need to represent this coupling honestly, likely through iterative modeling: geometry determines causal reach; matter distribution determines Φ sources; Φ (if tied to energy dissipation) may correlate with matter dynamics, feeding back into geometry.

11.3 Connections to causal set ideas, coarse-graining, and computational mechanics

The entire program of “causal topography” sits near a family of ideas that treat causality as fundamental and geometry as derived or secondary. Even without adopting any particular foundational stance, there are productive connections to explore.

First, causal set thinking suggests a shift of emphasis: instead of starting with a smooth manifold and adding Φ , start with a partially ordered set of events (a causal order) and define Φ on that order. In that setting, Model A becomes especially natural: “past-cone overlap with a record domain” becomes “count of elements in the causal past that lie in a designated subset.” In discrete form, measure becomes counting. This is conceptually clean and computationally accessible. It also clarifies what is essential: the partial order (what can influence what) plus a chosen subset representing record-bearing events.

Second, coarse-graining is not an embarrassment in this program; it is a central actor. Entropy and information are coarse-grained concepts. A promising future direction is to make coarse-graining explicit in the diagram by defining Φ at multiple scales:

- Φ_{micro} : tied to microscopic dynamics or local fluctuations
- Φ_{meso} : tied to mesoscopic structures like instruments, communication links, and memory substrates

- Φ_{macro} : tied to thermodynamic production and robust record formation

Then one can ask: how do these levels relate? Where do small-scale influences survive coarse-graining to become macroscopic records? Where do they wash out? This turns the diagram into a bridge between microphysics and the arrow of time, without pretending to solve foundational debates.

Third, computational mechanics offers a useful lens: it distinguishes between mere randomness and structured predictability, and it treats “stored information” and “predictive states” as physically meaningful. In that language, a record domain is not just “where matter is,” but “where predictive structure is maintained.” A future Φ might measure not just entropy production, but the creation and maintenance of predictive structure (or its loss). This is attractive because it aligns with the lived arrow: what matters to agents is not raw entropy but usable constraints—patterns that persist and can be exploited for prediction and control.

One open problem, therefore, is to define a Φ that distinguishes:

- high entropy that is featureless and not record-bearing
- high entropy that still contains structured constraints (redundant records, correlations)
- active entropy production that generates robust macroscopic marks
- passive equilibrium that is arrow-quiet even if entropy is high

This is not trivial, but it is exactly the kind of definition that would make causal topography more than “entropy as height.” It would become “history as structured constraint.”

Another connection is to the study of “information flow” measures in dynamical systems (influence in the sense of directed dependence). Many such measures depend on chosen variables and coarse-grainings, which fits the diagram’s honest stance. The open task is to select a small number of operationally interpretable flow measures that can be visualized as Φ without producing interpretive chaos.

11.4 Toward a standardized “causal topography” toolkit

If this program is to mature, it needs tooling—not just code, but conceptual interfaces. A standardized toolkit would not decree a single correct Φ . Instead it would provide a modular framework with clear choices, defaults, and diagnostics. The goal would be to make it easy for a researcher or teacher to say: “Here is the causal skeleton; here is my Φ definition; here is what it means; here are the failure modes.”

A plausible toolkit roadmap could include:

1. A library of Φ primitives
 - Accessibility primitives (past-cone overlap with domains, weighted by density and decay)
 - Propagation primitives (retarded kernels with configurable attenuation, dispersion, and noise)
 - Irreversibility primitives (entropy production proxies, localized macro-amplification events, dissipation fields)
 - Record robustness primitives (redundancy proxies, persistence weights, fragility maps)
2. Domain constructors for D
 - Static world-tubes, moving tubes, expanding/contracting tubes
 - Multiple domains with labels and weights
 - Domains derived from data (sensor coverage maps, communication graphs embedded in spacetime, matter density fields)
3. Invariance and frame-handling tools
 - Clear labeling of whether Φ is covariant scalar, frame-tied, or protocol-defined
 - Built-in transformations to show how a frame-tied Φ changes under boosts (to teach the distinction rather than hide it)
 - Optional use of proper time along specified worldlines for rate-based Φ
4. Visualization conventions
 - 1+1 lifted surface plots (the pedagogical “front end”)
 - 2+1 and 3+1 slice-based renderings
 - isosurfaces and causal-front overlays
 - standardized color/contour conventions that do not blur the line between geometry and field
5. Diagnostics and validation checks
 - Causal support check: ensure kernels vanish outside causal past/future regions
 - Robustness check: test sensitivity to grid resolution and smoothing

- Decomposition check: show component fields separately to avoid interpretive overload
- Interpretability check: a “legend generator” that outputs a human-readable statement of what Φ means and what it does not mean

6. Benchmark scenarios

A toolkit should ship with canonical scenarios that let users develop intuition and test definitions: a single impulse source, multiple sources, a moving observer, a record tube with decay, a localized dissipation event, a horizon-like exclusion, competing observer-labs, etc. Benchmarking is essential because otherwise each new Φ becomes incomparable to others.

The hardest open problem here is social rather than mathematical: standardization requires restraint. The toolkit must resist the temptation to become an everything-diagram. It should encourage users to start with one clean Φ definition and one clean question. In that sense, the most valuable feature of a causal-topography toolkit might be the constraints it enforces: “state your domain, state your measure, state your kernel, state your operational meaning.”

If successful, the outcome would be a new diagrammatic idiom: causal topography as a disciplined way to talk about “how the world writes history inside a symmetric causal skeleton.” It would remain compatible with relativity, compatible with thermodynamics, and flexible enough to support different levels of modeling—from toy pedagogical examples to data-informed influence maps—without ever confusing the base geometry with the field that lives on it.

In summary, the future directions cluster around three expansions: dimensional (3+1 visualization), geometric (curved spacetime and horizons), and semantic (better Φ definitions tied to records, predictability, and irreversible amplification). The unifying open problem is to keep the framework both expressive and honest: expressive enough to show arrow-structure, honest enough to never mistake arrow-structure for spacetime itself.

12. Conclusion: A New Visual Grammar for Causality and the Arrow

The standard Minkowski diagram is one of the rare visual tools in physics that can change the way a mind thinks. It compresses an entire causal logic into a picture. But it is also intentionally silent about the most human part of time: why histories feel one-way, why records accumulate, why the past is accessible while the future is not, and why “what actually happens” seems to carve a groove through a space of symmetric possibilities. The contribution of this paper is to propose a restrained extension—a third-axis field $\Phi(ct, x)$ that lives on the Minkowski plane—so that the diagram can show not only causal possibility but causal texture. The aim is not to add

metaphysics to relativity. The aim is to add a visual grammar that makes the arrow discussable in the same diagram that already makes causality discussable.

The deeper claim is pedagogical and methodological. If we can separate the invariant causal skeleton from the contingent arrow-content in a single, unified picture, we reduce a perennial confusion: the confusion between the symmetry of spacetime kinematics and the asymmetry of physical histories. The lifted diagram does not pretend the arrow is “in time.” It makes the arrow visible as a physically meaningful overlay—an emergent structure written by matter, boundary conditions, and record formation within the causal constraints of relativity.

12.1 Minkowski provides the skeleton; Φ provides lived asymmetry

The guiding discipline of the paper can be stated in one sentence: keep Minkowski intact, and place arrow-relevant structure above it rather than inside it.

Minkowski spacetime provides the skeleton: invariant intervals, light cones, and a strict boundary between what can and cannot influence an event. That skeleton is geometric and, in its kinematic form, time-reversal symmetric. It is the right substrate for any honest causal reasoning because it forbids “influence by rhetoric.” If something is spacelike separated, it is not causally connected. If something is inside the past cone, it could have mattered. This is the clean, hard logic of relativity.

Φ provides lived asymmetry precisely because lived asymmetry is not a property of the skeleton. It is a property of how the skeleton is occupied. Records do not exist everywhere. Signals do not propagate with equal strength everywhere. Irreversibility is not uniformly distributed. Entropy production is concentrated in processes that amplify microscopic differences into macroscopic marks. Memory and archives are physical. They live in world-tubes. They degrade. They replicate. They form redundantly, and that redundancy is what makes “the past” feel fixed.

When Φ is defined as causal accessibility (Model A), the diagram visualizes how much of a designated record domain lies in an event’s causal past. When Φ is defined as retarded influence propagation (Model B), it visualizes how influence actually arrives in fronts and wakes rather than merely being permitted. When Φ is defined as entropy production or irreversible amplification (Model C), it visualizes where the universe is actively hardening history—where “what happened” becomes embedded in many degrees of freedom. In each case, the Minkowski plane remains the reference, but the surface above it becomes a map of the world’s asymmetry as it is experienced by agents and embodied in matter.

The conclusion is not that one Φ is “the truth.” The conclusion is that the decomposition is the truth: skeleton below, arrow-content above. That decomposition is what standard diagrams cannot easily show.

12.2 The arrow as a field over spacetime, not a coordinate of spacetime

A major conceptual gain of the lifted diagram is that it makes a precise distinction that is often blurred in language: the arrow of time is not a direction added to the coordinate system; it is a pattern in physical processes that can be represented as a field.

Coordinates label events. Fields describe properties of events. Time is a coordinate in relativity. The arrow is not an extra coordinate. It is not a new axis in the metric. It is not a hidden dimension. It is a physically contingent asymmetry: an uneven distribution of records, irreversible amplification, and accessible information. That asymmetry can be mapped, just as temperature or density can be mapped, without implying that “temperature is a dimension of space.”

This matters because it dissolves a false dilemma. One is often pushed to choose between two extremes: either time’s arrow is “built into spacetime” (and therefore fundamental), or it is merely “psychological” (and therefore unreal). The field view offers a third stance: the arrow is real, but it is not kinematic. It is real as a distribution of physical conditions and processes over spacetime, constrained by causal structure but not reducible to it.

In the lifted diagram, the light cone remains symmetric, but Φ need not be. That is exactly how the world behaves. The laws may allow time reversal in a formal sense, but the realized history—given boundary conditions, matter content, and coarse-graining—has an orientation. The field depiction makes that orientation visible without committing the category error of treating it as a geometric primitive. It also leaves room for different levels of objectivity: some Φ definitions can be covariant scalars; others are properly observer-relative because they describe operational access to records. The diagram can hold both, as long as the distinction is made explicit.

The arrow-as-field view is therefore not merely a visualization trick. It is a conceptual hygiene practice: it forces the reader to ask, “Is this asymmetry in the geometry, or in the physical structure that sits on the geometry?” In nearly all macroscopic arrow discussions, the honest answer is the second.

12.3 Why this may sharpen intuition about records, memory, and irreversibility

The practical value of a new visual grammar is measured by the intuitions it makes easier.

First, it sharpens the intuition that “the past” is operationally “the recorded past.” Minkowski gives you the past cone, but the lifted diagram lets you see that only some parts of that cone are record-bearing and accessible. This turns a philosophical slogan into a geometric reading: high Φ regions correspond to rich potential archives; low Φ deserts correspond to causal poverty. The reader learns to separate causal reach from informational reach.

Second, it sharpens the intuition that irreversibility is local. It is tempting to speak as if entropy increases everywhere equally, but real irreversibility is produced by specific processes: dissipation, mixing, decoherence-like spreading into environments, and macroscopic amplification events. In the lifted diagram these become bumps, ramps, and cliffs—visible sites where the arrow is strong. That visibility helps the reader understand why “what happened” becomes hard: because the world paid the thermodynamic price to copy it into many degrees of freedom.

Third, it sharpens the intuition that causality has texture. Standard Minkowski diagrams are binary at the level of causal permission; the lifted diagram is graded. That grading aligns more closely with how influence works in practice: signals attenuate, channels fail, records degrade, and multiple sources overlap. The diagram becomes a causal weather map rather than a causal fence.

Finally, it offers a disciplined way to generate questions. Once you can draw Φ , you can ask where the ridges should be, where they should saturate, how they change with moving observers, how multiple record domains compete, and how horizons or curvature alter accessibility. Even when the answers are only at the toy-model level, the questions themselves are sharpened. The diagram gives you a vocabulary for thinking.

The concluding claim, then, is modest but consequential: by lifting the Minkowski diagram with a non-spatial field Φ , we gain a way to see in one frame what is usually split across two conceptual worlds. Relativity gives causal structure. Thermodynamics and information give arrow-structure. The lifted diagram keeps both visible without conflating them. That is the proposed new grammar: causality as skeleton, history as topography, and the arrow as a real, physically instantiated field written across spacetime rather than a hidden coordinate embedded within it.

References

1. Albert Einstein. Zur Elektrodynamik bewegter Körper. *Annalen der Physik* **322**(10), 891–921 (1905).
2. Einstein, A. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik* **354**(7), 769–822 (1916).
3. Hermann Minkowski. Raum und Zeit. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **18**, 75–88 (1909).
4. Edwin F. Taylor and John Archibald Wheeler. *Spacetime Physics* (2nd ed.) (1992).
5. Wolfgang Rindler. *Essential Relativity* (2nd ed.) (1977).
6. Robert M. Wald. *General Relativity* (1984).
7. Stephen Hawking and George F. R. Ellis. *The Large Scale Structure of Space-Time* (1973).
8. Roger Penrose. *Techniques of Differential Topology in Relativity*. *CBMS-NSF Regional Conference Series in Applied Mathematics* (1972).
9. Kronheimer, E. H., and Penrose, R. On the structure of causal spaces. *Proceedings of the Cambridge Philosophical Society* **63**, 481–501 (1967).
10. David B. Malament. The class of continuous timelike curves determines the topology of spacetime. *Journal of Mathematical Physics* **18**, 1399–1404 (1977).
11. Hawking, S. W., King, A. R., and McCarthy, P. J. A new topology for curved spacetime which incorporates the causal, differential, and conformal structures. *Journal of Mathematical Physics* **17**, 174–181 (1976).
12. Hans Reichenbach. *The Direction of Time* (1956).
13. H. Dieter Zeh. *The Physical Basis of the Direction of Time* (various editions).
14. Huw Price. *Time's Arrow and Archimedes' Point* (1996).
15. Sean Carroll. *From Eternity to Here* (2010).
16. Joel L. Lebowitz. Boltzmann's entropy and time's arrow. *Physics Today* **46**(9), 32–38 (1993).
17. de Groot, S. R., and Mazur, P. *Non-Equilibrium Thermodynamics* (1962).
18. Lars Onsager. Reciprocal relations in irreversible processes. I. *Physical Review* **37**, 405–426 (1931).

19. Ilya Prigogine. Introduction to Thermodynamics of Irreversible Processes (3rd ed.) (1967).
20. Israel, W., and Stewart, J. M. Transient relativistic thermodynamics and kinetic theory. *Annals of Physics* **118**(2), 341–372 (1979).
21. E. T. Jaynes. Information theory and statistical mechanics. *Physical Review* **106**, 620–630 (1957); and *Physical Review* **108**, 171–190 (1957).
22. Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–423 and 623–656 (1948).
23. Thomas M. Cover and Joy A. Thomas. Elements of Information Theory (2nd ed.) (2006).
24. Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* **5**(3), 183–191 (1961).
25. Charles H. Bennett. The thermodynamics of computation—A review. *International Journal of Theoretical Physics* **21**(12), 905–940 (1982).
26. Wojciech H. Zurek. Decoherence, einselection, and the quantum origins of the classical. *Reviews of Modern Physics* **75**, 715–775 (2003).
27. Maximilian Schlosshauer. Decoherence and the Quantum-to-Classical Transition (2007).
28. Joos, E., Zeh, H. D., Kiefer, C., Giulini, D., Kupsch, J., and Stamatescu, I.-O. Decoherence and the Appearance of a Classical World in Quantum Theory (2nd ed.) (2003).
29. Gell-Mann, M., and Hartle, J. B. Classical equations for quantum systems. *Physical Review D* **47**(8), 3345–3382 (1993).
30. Griffiths, R. B. Consistent histories and the interpretation of quantum mechanics. *Journal of Statistical Physics* **36**, 219–272 (1984).
31. Omnès, R. Consistent interpretations of quantum mechanics. *Reviews of Modern Physics* **64**, 339–382 (1992).
32. Thomas Schreiber. Measuring information transfer. *Physical Review Letters* **85**(2), 461–464 (2000).
33. James P. Crutchfield and Karl Young. Inferring statistical complexity. *Physical Review Letters* **63**, 105–108 (1989).
34. Shalizi, C. R., and Crutchfield, J. P. Computational mechanics: Pattern and prediction, structure and simplicity. *Journal of Statistical Physics* **104**, 817–879 (2001).

35. Bombelli, L., Lee, J., Meyer, D., and Sorkin, R. D. Space-time as a causal set. *Physical Review Letters* **59**, 521–524 (1987).
36. Rafael D. Sorkin. Causal Sets: Discrete Gravity (Notes for the Valdivia Summer School). arXiv:gr-qc/0309009 (2003).
37. Jacob D. Bekenstein. Black holes and entropy. *Physical Review D* **7**(8), 2333–2346 (1973).
38. Bekenstein, J. D. Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D* **23**, 287–298 (1981).
39. Raphael Bousso. A covariant entropy conjecture. *Journal of High Energy Physics* **1999**(07), 004 (1999).
40. Bousso, R. The holographic principle. *Reviews of Modern Physics* **74**(3), 825–874 (2002).

The author has published ~ 5,000 AI-assisted papers at:

<https://www.researchgate.net/profile/Douglas-Youvan/research> . Google Scholar has indexed these preprints, and if you want a particular reference, the AI mode of a Google search (Gemini) has efficiently catalogued these preprints.